crff 163903

Jarrell, Noble

From: Ramirez, Delia
Sent: Thursday, August 25, 2005 3:12 PM
To: Jarrell, Noble
Subject: FW: 09/598,982

-----Original Message-----
From: Ramirez, Delia
Sent: Monday, August 01, 2005 6:14 PM
To: Schreiber, David
Subject: 09/598,982

Hi,

I would like to request the following interference searches:

1. SEQ ID NO:52 in the protein and nucleic acid databases (commercial and interference)
2. SEQ ID NO:20 and 21 in the nucleic acid databases (interference)
3. SEQ ID NO:21 in the protein databases (interference)
4. an alignment of SEQ ID 9, 11, 21, 23, 25, 27, 37, 39, 41 and 43
5. an alignment of SEQ ID NO: 8, 10, 20, 22, 24, 26, 36, 38, 40, 42

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

Noble 8/31/05
Compugen 10PR
3NA 2AA 10 NA
Commercial
Interfer.

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

  http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

```
                     GenCore version 5.1.6
            Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         August 27, 2005, 21:12:24 ; Search time 607 Seconds
                                              (without alignments)
                                              2684.068 Million cell updates/sec

Title:          US-09-598-982C-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters:      14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_25082005_165520_304/app_query.fasta_1.391
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09598982_@CGN_1_1_723_@runat_25082005_165520_304
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
                19:  /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
                20:  /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
                21:  /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
                22:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                23:  /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
                24:  /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
                25:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                26:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                          Description
-----------------------------------------------------------------------------
     1    1368   98.2   1143  17  US-10-352-684A-47            Sequence 47, Appl
     2    1368   98.2   1145  18  US-10-287-226-93             Sequence 93, Appl
     3    1368   98.2   1148  21  US-10-956-157-2444           Sequence 2444, Ap
     4    1363   97.8    828  18  US-10-287-226-91             Sequence 91, Appl
     5    1363   97.8   1193  21  US-10-956-157-1043           Sequence 1043, Ap
     6    1355   97.3   1334  20  US-10-723-860-6799           Sequence 6799, Ap
     7    1344   96.5   1081  9  US-09-954-456-2126           Sequence 2126, Ap
     8    1344   96.5   1081  10  US-09-960-706-680            Sequence 680, App
     9    1344   96.5   1081  10  US-09-873-319-427            Sequence 427, App
    10    1344   96.5   1081  10  US-09-873-367C-155           Sequence 155, App
    11    1344   96.5   1081  10  US-09-873-367C-714           Sequence 714, App
    12    1344   96.5   1081  21  US-10-843-641A-155           Sequence 155, App
    13    1344   96.5   1081  21  US-10-843-641A-714           Sequence 714, App
    14    1344   96.5   1081  21  US-10-843-641A-5153          Sequence 5153, Ap
    15    1329   95.4   2662  18  US-10-275-505-27             Sequence 27, Appl
    16    1323   95.0    828  18  US-10-287-226-89             Sequence 89, Appl
    17    1312   94.2    858  18  US-10-287-226-87             Sequence 87, Appl
    18    1274   91.5   1154  17  US-10-352-684A-45            Sequence 45, Appl
    19    1274   91.5   1158  14  US-10-116-802-240            Sequence 240, App
    20    1274   91.5   1158  21  US-10-956-157-1042           Sequence 1042, Ap
    21    1232   88.4   5456  20  US-10-723-860-2689           Sequence 2689, Ap
    22    1232   88.4   5456  22  US-10-756-149-2484           Sequence 2484, Ap
    23    1054   75.7   1187  18  US-10-275-505-22             Sequence 22, Appl
    24    1005   72.1    729  21  US-10-480-988-55             Sequence 55, Appl
    25     988   70.9    729  21  US-10-956-157-1789           Sequence 1789, Ap
    26     958   68.8    600  21  US-10-956-157-7024           Sequence 7024, Ap
    27   664.5   47.7    846  10  US-09-813-432-11             Sequence 11, Appl
    28   664.5   47.7    846  17  US-10-174-364-11             Sequence 11, Appl
    29   664.5   47.7    846  18  US-10-246-583-11             Sequence 11, Appl
    30   664.5   47.7    846  19  US-10-689-832-11             Sequence 11, Appl
    31     634   45.5    873  14  US-10-117-323-2              Sequence 2, Appli
    32   605.5   43.5   1122  9  US-09-900-754-1              Sequence 1, Appli
    33   599.5   43.0    948  16  US-10-131-409-15             Sequence 15, Appl
    34   599.5   43.0    948  17  US-10-139-854-15             Sequence 15, Appl
    35   599.5   43.0    948  17  US-10-150-813-15             Sequence 15, Appl
    36   599.5   43.0    948  17  US-10-150-811-15             Sequence 15, Appl
    37   593.5   42.6   1199  21  US-10-956-157-1816           Sequence 1816, Ap
    38     584   41.9    978  15  US-10-311-955-1              Sequence 1, Appli
    39   583.5   41.9   1958  17  US-10-311-035-29             Sequence 29, Appl
    40   569.5   40.9    696  21  US-10-956-157-5078           Sequence 5078, Ap
    41   569.5   40.9    867  9  US-09-888-615-56             Sequence 56, Appl
    42   569.5   40.9    895  14  US-10-117-323-34             Sequence 34, Appl
    43   569.5   40.9   1218  17  US-10-274-639-36             Sequence 36, Appl
    44   569.5   40.9   1218  18  US-10-333-574-36             Sequence 36, Appl
    45     568   40.8    858  16  US-10-131-409-21             Sequence 21, Appl


                               ALIGNMENTS


RESULT 1
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
```

```
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
;   LENGTH: 1143
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6)...(833)
US-10-352-684A-47

Alignment Scores:
Pred. No.:              2.05e-158      Length:        1143
Score:                  1368.00        Matches:       244
Percent Similarity:     99.59%         Conservative:  0
Best Local Similarity:  99.59%         Mismatches:    1
Query Match:            98.21%         Indels:        0
DB:                     17             Gaps:          0

US-09-598-982C-21 (1-249) x US-10-352-684A-47 (1-1143)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 155

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 215

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 275

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 335

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 395

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 455

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 515

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 575

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 635

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 695

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 755

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        756 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 815

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        816 GTCCCCAAAAAGCCG 830

RESULT 2
US-10-287-226-93
; Sequence 93, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
```

```
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 93
;   LENGTH: 1145
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(832)
US-10-287-226-93

Alignment Scores:
Pred. No.:              2.06e-158     Length:        1145
Score:                  1368.00       Matches:       244
Percent Similarity:     99.59%        Conservative:  0
Best Local Similarity:  99.59%        Mismatches:    1
Query Match:            98.21%        Indels:        0
DB:                     18            Gaps:          0

US-09-598-982C-21 (1-249) x US-10-287-226-93 (1-1145)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 157

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 217

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 277

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 337

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 397

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 457

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 517

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 577

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 637

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 697

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 757

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 817

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        818 GTCCCCAAAAAGCCG 832

RESULT 3
US-10-956-157-2444
; Sequence 2444, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2444
;   LENGTH: 1148
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-2444

Alignment Scores:
Pred. No.:              2.06e-158     Length:        1148
Score:                  1368.00       Matches:       244
Percent Similarity:     99.59%        Conservative:  0
Best Local Similarity:  99.59%        Mismatches:    1
Query Match:            98.21%        Indels:        0
DB:                     21            Gaps:          0

US-09-598-982C-21 (1-249) x US-10-956-157-2444 (1-1148)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 157

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 217

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 277

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 337

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 397

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 457

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        458 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 517

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 577

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 637

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGCCCCTGGTGTGC 697

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 757

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 817

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        818 GTCCCCAAAAAGCCG 832

RESULT 4
US-10-287-226-91
; Sequence 91, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 91
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-91

Alignment Scores:
Pred. No.:              5.54e-158      Length:        828
Score:                  1363.00        Matches:       243
Percent Similarity:     99.18%         Conservative:  0
Best Local Similarity:  99.18%         Mismatches:    2
Query Match:            97.85%         Indels:        0
DB:                     18             Gaps:          0

US-09-598-982C-21 (1-249) x US-10-287-226-91 (1-828)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 150

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 210

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 270

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 330

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 390

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 450

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 510

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 570
```

```
Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 630

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 690

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 750

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 810

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        811 GTCCCCAAAAAGCCG 825

RESULT 5
US-10-956-157-1043
; Sequence 1043, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1043
;   LENGTH: 1193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-1043

Alignment Scores:
Pred. No.:              9e-158         Length:        1193
Score:                  1363.00        Matches:       243
Percent Similarity:     99.18%         Conservative:  0
Best Local Similarity:  99.18%         Mismatches:    2
Query Match:            97.85%         Indels:        0
DB:                     21             Gaps:          0

US-09-598-982C-21 (1-249) x US-10-956-157-1043 (1-1193)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 202

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 262

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 322

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 382

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 442

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 502
```

```
Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 562

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 622

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 682

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 742

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 802

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 862

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        863 GTCCCCAAAAAGCCG 877

RESULT 6
US-10-723-860-6799
; Sequence 6799, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6799
;   LENGTH: 1334
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-6799

Alignment Scores:
Pred. No.:              1.02e-156      Length:        1334
Score:                  1355.00        Matches:       242
Percent Similarity:     98.78%         Conservative:  0
Best Local Similarity:  98.78%         Mismatches:    3
Query Match:            97.27%         Indels:        0
DB:                     20             Gaps:          0

US-09-598-982C-21 (1-249) x US-10-723-860-6799 (1-1334)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 202

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CGCGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 262

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 322
```

```
Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 382

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 442

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 502

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 562

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 622

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 AAATACCACTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 682

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 742

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 802

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 862

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        863 GTCCCCAAAAAGCCG 877
```

```
RESULT 7
US-09-954-456-2126
; Sequence 2126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-2126

Alignment Scores:
Pred. No.:              1.75e-155       Length:        1081
Score:                  1344.00         Matches:       241
Percent Similarity:     98.37%          Conservative:  0
Best Local Similarity:  98.37%          Mismatches:    4
Query Match:            96.48%          Indels:        0
DB:                     9               Gaps:          0

US-09-598-982C-21 (1-249) x US-09-954-456-2126 (1-1081)
```

```
Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802
```

```
RESULT 8
US-09-960-706-680
; Sequence 680, Application US/09960706
; Publication No. US20030134280A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
US-09-960-706-680

Alignment Scores:
Pred. No.:                1.75e-155      Length:        1081
Score:                    1344.00        Matches:       241
Percent Similarity:       98.37%         Conservative:  0
Best Local Similarity:    98.37%         Mismatches:    4
Query Match:              96.48%         Indels:        0
DB:                       10             Gaps:          0

US-09-598-982C-21 (1-249) x US-09-960-706-680 (1-1081)

Qy         5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy        25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy        45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
             |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db       188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy        65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy        85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy       105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy       125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy       145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy       165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy       185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy       205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy       225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy       245 ValProLysLysPro 249
             |||||||||||||||
Db       788 GTCCCCAAAAAGCCG 802

RESULT 9
US-09-873-319-427
; Sequence 427, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 427
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
US-09-873-319-427

Alignment Scores:
Pred. No.:                1.75e-155      Length:        1081
Score:                    1344.00        Matches:       241
Percent Similarity:       98.37%         Conservative:  0
Best Local Similarity:    98.37%         Mismatches:    4
Query Match:              96.48%         Indels:        0
DB:                       10             Gaps:          0

US-09-598-982C-21 (1-249) x US-09-873-319-427 (1-1081)

Qy         5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy        25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy        45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
             |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db       188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy        65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy        85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy       105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy       125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 10
US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-155

Alignment Scores:
Pred. No.:              1.75e-155     Length:         1081
Score:                  1344.00       Matches:        241
Percent Similarity:     98.37%        Conservative:   0
Best Local Similarity:  98.37%        Mismatches:     4
Query Match:            96.48%        Indels:         0
DB:                     10            Gaps:           0

US-09-598-982C-21 (1-249) x US-09-873-367C-155 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                          ||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 11
US-09-873-367C-714
; Sequence 714, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 714
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-714

Alignment Scores:
Pred. No.:              1.75e-155     Length:         1081
Score:                  1344.00       Matches:        241
Percent Similarity:     98.37%        Conservative:   0
Best Local Similarity:  98.37%        Mismatches:     4
Query Match:            96.48%        Indels:         0
DB:                     10            Gaps:           0

US-09-598-982C-21 (1-249) x US-09-873-367C-714 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 12
US-10-843-641A-155
; Sequence 155, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-155

Alignment Scores:
Pred. No.:              1.75e-155     Length:         1081
Score:                  1344.00       Matches:        241
Percent Similarity:     98.37%        Conservative:   0
Best Local Similarity:  98.37%        Mismatches:     4
Query Match:            96.48%        Indels:         0
DB:                     21            Gaps:           0

US-09-598-982C-21 (1-249) x US-10-843-641A-155 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487
```

```
Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 13
US-10-843-641A-714
; Sequence 714, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 714
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-714

Alignment Scores:
Pred. No.:              1.75e-155     Length:        1081
Score:                  1344.00       Matches:       241
Percent Similarity:     98.37%        Conservative:  0
Best Local Similarity:  98.37%        Mismatches:    4
Query Match:            96.48%        Indels:        0
DB:                     21            Gaps:          0

US-09-598-982C-21 (1-249) x US-10-843-641A-714 (1-1081)
```

```
Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 14
US-10-843-641A-5153
; Sequence 5153, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
```

```
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5153
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-5153

Alignment Scores:
Pred. No.:              1.75e-155      Length:        1081
Score:                  1344.00        Matches:       241
Percent Similarity:     98.37%         Conservative:  0
Best Local Similarity:  98.37%         Mismatches:    4
Query Match:            96.48%         Indels:        0
DB:                     21             Gaps:          0

US-09-598-982C-21 (1-249) x US-10-843-641A-5153 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727
```

```
Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 15
US-10-275-505-27
; Sequence 27, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 27
;   LENGTH: 2662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7472460CB1
US-10-275-505-27

Alignment Scores:
Pred. No.:              4.13e-153      Length:        2662
Score:                  1329.00        Matches:       238
Percent Similarity:     97.14%         Conservative:  0
Best Local Similarity:  97.14%         Mismatches:    7
Query Match:            95.41%         Indels:        0
DB:                     18             Gaps:          0

US-09-598-982C-21 (1-249) x US-10-275-505-27 (1-2662)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 760

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 820

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         821 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 880

Qy          65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         881 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 940

Qy          85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         941 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 1000

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
               |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1001 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 1060

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1061 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 1120

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
               ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        1121 CCGCCATTTCCTCTGAAGCATGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 1180

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1181 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 1240

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1241 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 1300

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
               ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        1301 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGATGGGGAGAGGGCTGTGCCCAG 1360

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1361 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 1420

Qy         245 ValProLysLysPro 249
               |||||||||||||||
Db        1421 GTCCCCAAAAAGCCG 1435

Search completed: August 27, 2005, 23:11:52
Job time : 613 secs
```

```
                          GenCore version 5.1.6
                 Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:          August 27, 2005, 19:35:09 ; Search time 178 Seconds
                                          (without alignments)
                                          2288.949 Million cell updates/sec

Title:           US-09-598-982C-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:      2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_25082005_165518_269/app_query.fasta_1.391
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_105_@runat_25082005_165518_269 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                       Description
-------------------------------------------------------------------------------
    1    1387    99.6    771  3  US-09-079-970A-4          Sequence 4, Appli
    2    1368    98.2    735  3  US-09-079-970A-1          Sequence 1, Appli
    3    1368    98.2   1128  2  US-09-016-366A-20         Sequence 20, Appl
    4    1368    98.2   1128  2  US-08-978-404B-15         Sequence 15, Appl
    5    1363    97.8   1137  2  US-09-016-366A-18         Sequence 18, Appl
    6    1363    97.8   1137  2  US-08-978-404B-13         Sequence 13, Appl
    7    1344    96.5   1081  2  US-09-016-366A-22         Sequence 22, Appl
    8    1344    96.5   1081  2  US-08-978-404B-17         Sequence 17, Appl
    9    1344    96.5   1081  4  US-09-917-254-50          Sequence 50, Appl
   10    1274    91.5   1154  2  US-09-016-366A-16         Sequence 16, Appl
   11    1274    91.5   1154  2  US-08-978-404B-11         Sequence 11, Appl
   12    1091    78.3   1108  2  US-09-016-366A-14         Sequence 14, Appl
   13    1091    78.3   1108  2  US-08-978-404B-20         Sequence 20, Appl
   14    1075    77.2   1219  2  US-08-978-404B-7          Sequence 7, Appli
   15    1054    75.7   1031  2  US-08-978-404B-1          Sequence 1, Appli
   16    1051    75.4   1097  2  US-08-978-404B-4          Sequence 4, Appli
   17    1051    75.4   1103  2  US-09-016-366A-24         Sequence 24, Appl
   18  1000.5    71.8   2259  2  US-08-845-998-3           Sequence 3, Appli
   19  1000.5    71.8   2259  3  US-09-206-537-3           Sequence 3, Appli
   20  1000.5    71.8   2259  3  US-09-430-854-3           Sequence 3, Appli
   21     979    70.3   3757  2  US-09-016-366A-13         Sequence 13, Appl
   22     979    70.3   3757  2  US-08-978-404B-19         Sequence 19, Appl
   23   961.5    69.0   2218  2  US-08-845-998-5           Sequence 5, Appli
   24   961.5    69.0   2218  3  US-09-206-537-5           Sequence 5, Appli
   25   961.5    69.0   2218  3  US-09-430-854-5           Sequence 5, Appli
   26   836.5    60.1   2397  2  US-08-978-404B-2          Sequence 2, Appli
   27   684.5    49.1   1095  2  US-08-978-404B-9          Sequence 9, Appli
   28   563.5    40.5    980  4  US-09-023-942A-30         Sequence 30, Appl
   29   563.5    40.5   1110  3  US-09-386-653A-1          Sequence 1, Appli
   30   561.5    40.3   1212  4  US-09-620-312D-431        Sequence 431, App
   31   554.5    39.8   1130  3  US-09-386-653A-8          Sequence 8, Appli
   32     552    39.6    933  4  US-09-023-942A-29         Sequence 29, Appl
   33   536.5    38.5   1613  4  US-09-387-375-1           Sequence 1, Appli
   34   536.5    38.5   1613  4  US-10-041-400A-1          Sequence 1, Appli
   35   536.5    38.5   1613  4  US-10-042-091A-1          Sequence 1, Appli
   36   527.5    37.9   1130  4  US-09-387-375-8           Sequence 8, Appli
   37   527.5    37.9   1130  4  US-10-041-400A-8          Sequence 8, Appli
   38   527.5    37.9   1130  4  US-10-042-091A-8          Sequence 8, Appli
   39     507    36.4   1142  3  US-09-386-642-8           Sequence 8, Appli
   40     507    36.4   1169  3  US-09-386-642-7           Sequence 7, Appli
   41   501.5    36.0   1165  4  US-09-023-942A-28         Sequence 28, Appl
   42   501.5    36.0   1378  4  US-09-907-794A-262        Sequence 262, App
   43   501.5    36.0   1378  4  US-09-905-125A-262        Sequence 262, App
   44   501.5    36.0   1378  4  US-09-902-775A-262        Sequence 262, App
   45   501.5    36.0   1378  4  US-09-906-700-262         Sequence 262, App


                               ALIGNMENTS

RESULT 1
US-09-079-970A-4
; Sequence 4, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;   INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  7..753
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7..18
US-09-079-970A-4

Alignment Scores:
Pred. No.:              3.11e-150     Length:        771
Score:                  1387.00       Matches:       248
Percent Similarity:     99.60%        Conservative:  0
Best Local Similarity:  99.60%        Mismatches:    1
Query Match:            99.57%        Indels:        0
DB:                     3             Gaps:          0

US-09-598-982C-21 (1-249) x US-09-079-970A-4 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 2
US-09-079-970A-1
; Sequence 1, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..735
US-09-079-970A-1

Alignment Scores:
Pred. No.:              4.49e-148     Length:        735
Score:                  1368.00       Matches:       244
Percent Similarity:     99.59%        Conservative:  0
Best Local Similarity:  99.59%        Mismatches:    1
Query Match:            98.21%        Indels:        0
DB:                     3             Gaps:          0

US-09-598-982C-21 (1-249) x US-09-079-970A-1 (1-735)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120
```

```
Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

RESULT 3
US-09-016-366A-20
; Sequence 20, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-20

Alignment Scores:
Pred. No.:                8.69e-148      Length:         1128
Score:                    1368.00        Matches:        244
Percent Similarity:       99.59%         Conservative:   0
Best Local Similarity:    99.59%         Mismatches:     1
Query Match:              98.21%         Indels:         0
DB:                       2              Gaps:           0

US-09-598-982C-21 (1-249) x US-09-016-366A-20 (1-1128)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 148

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 208

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 268

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 328

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 388

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 448

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 508

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 568

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 628

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 688

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 748

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          749 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 808

Qy          245 ValProLysLysPro 249
                |||||||||||||||
Db          809 GTCCCCAAAAAGCCG 823

RESULT 4
US-08-978-404B-15
; Sequence 15, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-15

Alignment Scores:
Pred. No.:               8.69e-148      Length:        1128
Score:                   1368.00        Matches:       244
Percent Similarity:      99.59%         Conservative:  0
Best Local Similarity:   99.59%         Mismatches:    1
Query Match:             98.21%         Indels:        0
DB:                      2              Gaps:          0

US-09-598-982C-21 (1-249) x US-08-978-404B-15 (1-1128)

Qy            5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           89 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 148

Qy           25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          149 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 208

Qy           45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db          209 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 268

Qy           65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          269 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 328

Qy           85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          329 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 388

Qy          105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          389 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 448

Qy          125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          449 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 508

Qy          145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          509 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 568

Qy          165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          569 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 628

Qy          185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          629 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 688

Qy          205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          689 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 748

Qy          225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          749 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 808

Qy          245 ValProLysLysPro 249
                |||||||||||||||
Db          809 GTCCCCAAAAAGCCG 823

RESULT 5
US-09-016-366A-18
; Sequence 18, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-18

Alignment Scores:
Pred. No.:              3.32e-147      Length:        1137
Score:                  1363.00        Matches:       243
Percent Similarity:     99.18%         Conservative:  0
Best Local Similarity:  99.18%         Mismatches:    2
Query Match:            97.85%         Indels:        0
DB:                     2              Gaps:          0

US-09-598-982C-21 (1-249) x US-09-016-366A-18 (1-1137)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 146

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 206

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 266

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 326

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 386

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 446

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 506

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 566

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 626

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 686

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 746

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 806

Qy        245 ValProLysLysPro 249
              |||||||||||||||
```

```
Db        807 GTCCCCAAAAAGCCG 821

RESULT 6
US-08-978-404B-13
; Sequence 13, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-13

Alignment Scores:
Pred. No.:              3.32e-147      Length:        1137
Score:                  1363.00        Matches:       243
Percent Similarity:     99.18%         Conservative:  0
Best Local Similarity:  99.18%         Mismatches:    2
Query Match:            97.85%         Indels:        0
DB:                     2              Gaps:          0

US-09-598-982C-21 (1-249) x US-08-978-404B-13 (1-1137)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 146

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 206

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 266

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 326

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           327 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 386

Qy           105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                 |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           387 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 446

Qy           125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           447 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 506

Qy           145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           507 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 566

Qy           165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           567 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 626

Qy           185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           627 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 686

Qy           205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           687 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 746

Qy           225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           747 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 806

Qy           245 ValProLysLysPro 249
                 |||||||||||||||
Db           807 GTCCCCAAAAAGCCG 821

RESULT 7
US-09-016-366A-22
; Sequence 22, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  22:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-22

Alignment Scores:
Pred. No.:              4.76e-145      Length:        1081
Score:                  1344.00        Matches:       241
Percent Similarity:     98.37%         Conservative:  0
Best Local Similarity:  98.37%         Mismatches:    4
Query Match:            96.48%         Indels:        0
DB:                     2              Gaps:          0

US-09-598-982C-21 (1-249) x US-09-016-366A-22 (1-1081)

Qy             5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy            25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy            45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                 |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db           188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy            65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy            85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy           105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy           125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy           145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy           165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy           185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy           205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy           225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy           245 ValProLysLysPro 249
                 |||||||||||||||
Db           788 GTCCCCAAAAAGCCG 802

RESULT 8
US-08-978-404B-17
; Sequence 17, Application US/08978404B
```

```
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-17

Alignment Scores:
Pred. No.:                4.76e-145      Length:        1081
Score:                    1344.00        Matches:       241
Percent Similarity:       98.37%         Conservative:  0
Best Local Similarity:    98.37%         Mismatches:    4
Query Match:              96.48%         Indels:        0
DB:                       2              Gaps:          0

US-09-598-982C-21 (1-249) x US-08-978-404B-17 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 9
US-09-917-254-50
; Sequence 50, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Baak, Jan
;  TITLE OF INVENTION: Prognostic Classification of Breast Cancer
;  FILE REFERENCE: B0801/7224(JRV)
;  CURRENT APPLICATION NUMBER: US/09/917,254
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/222,093
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-917-254-50

Alignment Scores:
Pred. No.:                4.76e-145      Length:        1081
Score:                    1344.00        Matches:       241
Percent Similarity:       98.37%         Conservative:  0
Best Local Similarity:    98.37%         Mismatches:    4
Query Match:              96.48%         Indels:        0
DB:                       4              Gaps:          0

US-09-598-982C-21 (1-249) x US-09-917-254-50 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247
```

```
Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 10
US-09-016-366A-16
; Sequence 16, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-16

Alignment Scores:
Pred. No.:              6.2e-137        Length:         1154
Score:                  1274.00         Matches:        226
Percent Similarity:     94.69%          Conservative:   6
Best Local Similarity:  92.24%          Mismatches:     13
Query Match:            91.46%          Indels:         0
DB:                     2               Gaps:           0

US-09-598-982C-21 (1-249) x US-09-016-366A-16 (1-1154)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 167

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 227

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        228 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 287

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCAC 347

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        348 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 407

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        408 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 467

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        468 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 527

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 587

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        588 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 647

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db        648 TGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGCCCCCTGGTGTGC 707

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        708 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCCAG 767

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 827

Qy        245 ValProLysLysPro 249
              |||||||||||||||
```

```
Db        828 GTCCCCAAAAAGCCG 842

RESULT 11
US-08-978-404B-11
; Sequence 11, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-11

Alignment Scores:
Pred. No.:              6.2e-137      Length:        1154
Score:                  1274.00       Matches:       226
Percent Similarity:     94.69%        Conservative:  6
Best Local Similarity:  92.24%        Mismatches:    13
Query Match:            91.46%        Indels:        0
DB:                     2             Gaps:          0

US-09-598-982C-21 (1-249) x US-08-978-404B-11 (1-1154)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 167

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 227

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        228 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 287

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCAC 347

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        348 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 407

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        408 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 467

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        468 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 527

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 587

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        588 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 647

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db        648 TGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGTGC 707

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        708 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCCAG 767

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 827

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        828 GTCCCCAAAAAGCCG 842

RESULT 12
US-09-016-366A-14
; Sequence 14, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-14

Alignment Scores:
Pred. No.:              7.31e-116      Length:         1108
Score:                  1091.00        Matches:        190
Percent Similarity:     86.01%         Conservative:   19
Best Local Similarity:  78.19%         Mismatches:     34
Query Match:            78.32%         Indels:         0
DB:                     2              Gaps:           0

US-09-598-982C-21 (1-249) x US-09-016-366A-14 (1-1108)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||   ||||||      ||||||||||||||||||||||||||||||
Db        127 ATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 186

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       ||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        187 AAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTC 246

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||   :::|||               ||||||||||||
Db        247 ACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTT 306

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||||||||:::|||||||||   ||||||||||||   ::::::||||||:::||||||
Db        307 CGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCAC 366

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   :::|||||||||:::   |||||||||:::||||||||||||||||||   |||
Db        367 CCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCT 426

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
Db        427 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 486

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||      ||||||||||||||||||||||||:::||||||||||||   ||||||
Db        487 CCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCA 546

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db        547 CCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGG 606

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||   |||||||||||||||      ||||||   |||   ||||||
Db        607 AAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTG 666

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCTGC 726

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 AAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCACAG 786

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||   |||
Db        787 CCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCTAT 846

Qy        245 ValProLys 247
              ||||||:::
Db        847 GTCCCTGAG 855

RESULT 13
US-08-978-404B-20
; Sequence 20, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-20

Alignment Scores:
Pred. No.:              7.31e-116      Length:         1108
Score:                  1091.00        Matches:        190
Percent Similarity:     86.01%         Conservative:   19
Best Local Similarity:  78.19%         Mismatches:     34
Query Match:            78.32%         Indels:         0
DB:                     2              Gaps:           0

US-09-598-982C-21 (1-249) x US-08-978-404B-20 (1-1108)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||   ||||||      ||||||||||||||||||||||||||||||
Db        127 ATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 186

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       ||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        187 AAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTC 246

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||   :::|||               ||||||||||||
Db        247 ACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTT 306

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||||||||:::|||||||||   ||||||||||||   ::::::||||||:::||||||
Db        307 CGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCAC 366

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   :::|||||||||:::   |||||||||:::||||||||||||||||||   |||
Db        367 CCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCT 426

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
```

```
Db         427 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 486

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
               ||||||      ||||||||||||||||||||||||:::||||||||||||   ||||||
Db         487 CCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCA 546

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
               ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db         547 CCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGG 606

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
               |||||||||   |||   |||||||||||||||      ||||||   |||   ||||||
Db         607 AAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTG 666

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         667 TGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCTGC 726

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         727 AAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCACAG 786

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
               ||||||:::|||||||||||||||||||||||||||||||||||||||||||||   |||
Db         787 CCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCTAT 846

Qy         245 ValProLys 247
               ||||||:::
Db         847 GTCCCTGAG 855

RESULT 14
US-08-978-404B-7
; Sequence 7, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1219 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-7

Alignment Scores:
Pred. No.:              5.93e-114      Length:        1219
Score:                  1075.00        Matches:       188
Percent Similarity:     84.77%         Conservative:  18
Best Local Similarity:  77.37%         Mismatches:    37
Query Match:            77.17%         Indels:        0
DB:                     2              Gaps:          0

US-09-598-982C-21 (1-249) x US-08-978-404B-7 (1-1219)

Qy           5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
               ||||||||||||||||||||||||   :::|||||||||||||||||||||||||||
Db         273 ATTGTTGGGGGACAGGAGGCACCTGGGAACAAGTGGCCCTGGCAGGTGAGCCTTCGTGCC 332

Qy          25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
               :::      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db         333 AATGAAACCTACTGGAGGCATTTCTGCGGCGGCTCCCTCATCCACCCACAGTGGGTGCTC 392

Qy          45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
               |||||||||   ||||||||||||   :::   |||         :::||||||||||||
Db         393 ACCGCGGCACACTGTGTGGGACCGACTATTGCTGATCCCAACAAGGTCAGAGTACAGCTT 452

Qy          65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
               |||:::|||:::|||||||||   |||   ||||||   |||||||||||||||   |||
Db         453 CGAAAGCAGTACCTCTATTACCACGACCACCTGCTGGCTGTGAGCCGGATCATCACACAC 512

Qy          85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
               |||   ||||||      |||   |||||||||||||||||||||||||||:::   |||
Db         513 CCGACATTCTATGCCACCCAGAATGGGGCGGACATCGCCCTACTTGAGCTCAAGAACCCT 572

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
               |||   :::|||||||||||||||   |||:::|||||||||||||||||||||||||||
Db         573 GTAAACATTTCCAGCCATGTCCACCCCGTCTCCCTGCCTCCTGCCTCAGAGACCTTCCCC 632

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                  |||      |||||||||||||||||||||::::::|||||||||      ||||||
Db         633 TCAGGAACATTGTGCTGGGTGACAGGCTGGGGAAACATCGACAATGATGTGAGCCTGCCA 692

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
               ||||||||||||||||||:::|||:::||||||::::::||||||   :::||||||
Db         693 CCGCCATTTCCCTTGAAGGAGGTGCAAGTTCCCGTCGTGGAAAACCAGCTTTGTGACCTG 752

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
               |||||||||   |||   ||||||||||||::::::   |||||||||||||||||||||
Db         753 AAGTATCACAAAGGTGTCTACACAGGGGACAACATCCACATTGTCCGAGACGACATGCTG 812

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
               ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db         813 TGTGCTGGGAACGAAGGACACGACTCCTGCCAGGGTGACTCCGGAGGACCTCTGGTCTGC 872

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         873 AAGGTAAACGGTACCTGGCTGCAGGCAGGTGTGGTCAGCTGGGGTGAGGGCTGTGCTCTG 932

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         933 CCCAACAGGCCTGGCATCTACACTCGGGTCACCTATTACTTGGACTGGATCCACCGCTAT 992

Qy         245 ValProLys 247
               |||||||||
Db         993 GTCCCCAAG 1001

RESULT 15
US-08-978-404B-1
; Sequence 1, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;       STREET:  600 Atlantic Avenue
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  U.S.A.
;       ZIP:  02210-2211
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/978,404B
;       FILING DATE:  25-NOV-97
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/032,354
;       FILING DATE:  04-DEC-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Plumer, Elizabeth R.
;       REGISTRATION NUMBER:  36,637
;       REFERENCE/DOCKET NUMBER:  B0801/7090
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-720-3500
;       TELEFAX:  617-720-2441
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1031 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-978-404B-1

Alignment Scores:
Pred. No.:              1.21e-111     Length:         1031
Score:                  1054.00       Matches:        186
Percent Similarity:     83.54%        Conservative:   17
Best Local Similarity:  76.54%        Mismatches:     40
Query Match:            75.66%        Indels:         0
DB:                     2             Gaps:           0

US-09-598-982C-21 (1-249) x US-08-978-404B-1 (1-1031)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              |||||||||||||||||||||      :::|||||||||||||||||||||||||||   
Db        111 ATTGTGGGGGGACAGGAGGCACATGGGAACAAGTGGCCCTGGCAGGTGAGCCTGCGTGCC 170

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              :::      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 AATGACACCTACTGGATGCATTTCTGCGGTGGCTCCCTCATCCACCCACAGTGGGTGCTC 230

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||   |||         :::||||||||||||
Db        231 ACTGCGGCACACTGTGTGGGACCGGATGTTGCTGACCCCAACAAGGTCAGAGTACAGCTC 290

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||:::|||:::|||||||||   |||   |||:::   ||||||:::||||||   |||
Db        291 CGTAAGCAGTACCTCTATTACCATGACCACCTGATGACTGTGAGCCAGATCATCACACAC 350

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   ||||||      |||   |||||||||||||||||||||:::|||      |||
Db        351 CCCGACTTCTACATCGTCCAGGATGGGGCAGACATTGCCCTGCTGAAACTCACAAACCCT 410

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::|||   :::||||||   |||   |||||||||||||||||||||||||||
Db        411 GTGAACATTTCTGACTATGTCCACCCTGTCCCCCTACCTCCTGCCTCAGAGACCTTCCCC 470

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 |||      |||||||||||||||||||||::::::||||||         ||||||
Db        471 TCAGGAACGTTGTGCTGGGTGACAGGCTGGGGTAACATCGACAATGGTGTAAACCTGCCG 530

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||:::|||:::|||||||||:::|||||||||:::||||||   
Db        531 CCACCATTTCCTTTGAAGGAGGTGCAAGTTCCCATTATAGAAAACCACCTTTGTGACTTG 590

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||      |||||||||:::|||   |||||||||||||||||||||
Db        591 AAGTATCACAAAGGTCTCATCACAGGTGACAATGTCCACATTGTCCGAGATGACATGCTG 650

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db        651 TGTGCTGGGAATGAAGGACATGACTCCTGCCAGGGCGACTCCGGAGGACCTCTGGTCTGC 710

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AAGGTAGAAGACACCTGGCTGCAGGCAGGCGTGGTCAGCTGGGGTGAGGGCTGTGCACAG 770

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 CCCAACAGGCCTGGCATCTACACCCGGGTCACCTATTACTTGGACTGGATCCACCACTAT 830

Qy        245 ValProLys 247
              |||||||||
Db        831 GTCCCCAAG 839

Search completed: August 27, 2005, 21:26:04
Job time : 183 secs
```

10

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 27, 2005, 23:01:40 ; Search time 113 Seconds
                                          (without alignments)
                                          865.829 Million cell updates/sec

Title:          US-09-598-982C-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters:      1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                16:  /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
                17:  /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
                18:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                19:  /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
                20:  /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
                21:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                22:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                       Description
 ---------------------------------------------------------------------------
    1    1368    98.2    254  14  US-10-131-409-74         Sequence 74, Appl
    2    1368    98.2    254  15  US-10-139-854-74         Sequence 74, Appl
    3    1368    98.2    254  15  US-10-150-813-74         Sequence 74, Appl
    4    1368    98.2    254  15  US-10-150-811-74         Sequence 74, Appl
    5    1368    98.2    256  14  US-10-131-409-92         Sequence 92, Appl
    6    1368    98.2    256  15  US-10-139-854-92         Sequence 92, Appl
    7    1368    98.2    256  15  US-10-150-813-92         Sequence 92, Appl
    8    1368    98.2    256  15  US-10-150-811-92         Sequence 92, Appl
    9    1368    98.2    264  14  US-10-131-409-104        Sequence 104, App
   10    1368    98.2    264  15  US-10-139-854-104        Sequence 104, App
   11    1368    98.2    264  15  US-10-150-813-104        Sequence 104, App
   12    1368    98.2    264  15  US-10-150-811-104        Sequence 104, App
   13    1368    98.2    275  10  US-09-813-432-44         Sequence 44, Appl
   14    1368    98.2    275  15  US-10-352-684A-48        Sequence 48, Appl
   15    1368    98.2    275  15  US-10-174-364-44         Sequence 44, Appl
   16    1368    98.2    275  15  US-10-246-583-44         Sequence 44, Appl
   17    1368    98.2    275  15  US-10-287-226-94         Sequence 94, Appl
   18    1368    98.2    275  16  US-10-689-832-44         Sequence 44, Appl
   19    1363    97.8    275  15  US-10-287-226-92         Sequence 92, Appl
   20    1363    97.8    275  16  US-10-723-860-2690       Sequence 2690, Ap
   21    1363    97.8    275  18  US-10-756-149-5451       Sequence 5451, Ap
   22  1347.5    96.7    272  14  US-10-117-323-40         Sequence 40, Appl
   23    1329    95.4    691  15  US-10-275-505-13         Sequence 13, Appl
   24    1323    95.0    275  15  US-10-287-226-90         Sequence 90, Appl
   25    1317    94.5    266  14  US-10-131-409-84         Sequence 84, Appl
   26    1317    94.5    266  15  US-10-139-854-84         Sequence 84, Appl
   27    1317    94.5    266  15  US-10-150-813-84         Sequence 84, Appl
   28    1317    94.5    266  15  US-10-150-811-84         Sequence 84, Appl
   29    1312    94.2    279  15  US-10-287-226-88         Sequence 88, Appl
   30    1274    91.5    245  11  US-09-789-210-69         Sequence 69, Appl
   31    1274    91.5    275  14  US-10-117-323-36         Sequence 36, Appl
   32    1274    91.5    275  15  US-10-352-684A-46        Sequence 46, Appl
   33    1270    91.2    231  10  US-09-898-837A-52        Sequence 52, Appl
   34    1104    79.3    199  14  US-10-131-409-98         Sequence 98, Appl
   35    1104    79.3    199  15  US-10-139-854-98         Sequence 98, Appl
   36    1104    79.3    199  15  US-10-150-813-98         Sequence 98, Appl
   37    1104    79.3    199  15  US-10-150-811-98         Sequence 98, Appl
   38    1091    78.3    252  10  US-09-813-432-42         Sequence 42, Appl
   39    1091    78.3    252  15  US-10-174-364-42         Sequence 42, Appl
   40    1091    78.3    252  15  US-10-246-583-42         Sequence 42, Appl
   41    1091    78.3    252  16  US-10-689-832-42         Sequence 42, Appl
   42    1091    78.3    276  14  US-10-117-323-35         Sequence 35, Appl
   43    1077    77.3    276  14  US-10-311-955-3          Sequence 3, Appli
   44    1064    76.4    247  14  US-10-131-409-86         Sequence 86, Appl
   45    1064    76.4    247  15  US-10-139-854-86         Sequence 86, Appl

                              ALIGNMENTS

RESULT 1
US-10-131-409-74
; Sequence 74, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON1
;  CURRENT APPLICATION NUMBER: US/10/131,409
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
```

```
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-409-74

  Query Match             98.2%;  Score 1368;  DB 14;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy        245 VPKKP 249
              |||||
Db        250 VPKKP 254

RESULT 2
US-10-139-854-74
; Sequence 74, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder, Kumud
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CON2
;  CURRENT APPLICATION NUMBER: US/10/139,854
;  CURRENT FILING DATE:  2002-12-02
;  PRIOR APPLICATION NUMBER: 09/783,429
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-139-854-74

  Query Match             98.2%;  Score 1368;  DB 15;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy        245 VPKKP 249
              |||||
Db        250 VPKKP 254

RESULT 3
US-10-150-813-74
; Sequence 74, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON2
;  CURRENT APPLICATION NUMBER: US/10/150,813
;  CURRENT FILING DATE:  2002-12-03
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-813-74

  Query Match             98.2%;  Score 1368;  DB 15;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy        245 VPKKP 249
              |||||
Db        250 VPKKP 254


RESULT 4
US-10-150-811-74
; Sequence 74, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-675CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/150,811
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 09/970,607
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-811-74

  Query Match             98.2%;  Score 1368;  DB 15;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 4.2e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy        245 VPKKP 249
              |||||
Db        250 VPKKP 254


RESULT 5
US-10-131-409-92
; Sequence 92, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON1
;  CURRENT APPLICATION NUMBER: US/10/131,409
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-409-92

  Query Match             98.2%;  Score 1368;  DB 14;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256


RESULT 6
US-10-139-854-92
; Sequence 92, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder, Kumud
```

```
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CON2
;  CURRENT APPLICATION NUMBER: US/10/139,854
;  CURRENT FILING DATE:  2002-12-02
;  PRIOR APPLICATION NUMBER: 09/783,429
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-139-854-92

  Query Match             98.2%;  Score 1368;  DB 15;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256

RESULT 7
US-10-150-813-92
; Sequence 92, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON2
;  CURRENT APPLICATION NUMBER: US/10/150,813
;  CURRENT FILING DATE:  2002-12-03
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-813-92

  Query Match             98.2%;  Score 1368;  DB 15;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256

RESULT 8
US-10-150-811-92
; Sequence 92, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/150,811
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 09/970,607
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
```

```
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-811-92

  Query Match             98.2%;  Score 1368;  DB 15;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 4.3e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256

RESULT 9
US-10-131-409-104
; Sequence 104, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-675CIP1CON1
;  CURRENT APPLICATION NUMBER: US/10/131,409
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-409-104

  Query Match             98.2%;  Score 1368;  DB 14;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264

RESULT 10
US-10-139-854-104
; Sequence 104, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder, Kumud
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CON2
;  CURRENT APPLICATION NUMBER: US/10/139,854
;  CURRENT FILING DATE:  2002-12-02
;  PRIOR APPLICATION NUMBER: 09/783,429
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-139-854-104

  Query Match             98.2%;  Score 1368;  DB 15;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264


RESULT 11
US-10-150-813-104
; Sequence 104, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON2
;  CURRENT APPLICATION NUMBER: US/10/150,813
;  CURRENT FILING DATE:  2002-12-03
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-813-104

  Query Match             98.2%;  Score 1368;  DB 15;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264


RESULT 12
US-10-150-811-104
; Sequence 104, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/150,811
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 09/970,607
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-811-104

  Query Match             98.2%;  Score 1368;  DB 15;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
```

```
                   |||||
Db           260 VPKKP 264

RESULT 13
US-09-813-432-44
; Sequence 44, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier Jr., Raymond J
;  APPLICANT:  Majmuder, Kamud
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Smithson, Glenda
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Vernet, Corine A. M.
;  TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
;  FILE REFERENCE: 15966-729
;  CURRENT APPLICATION NUMBER: US/09/813,432
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 60/193,843
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-813-432-44

  Query Match             98.2%;  Score 1368;  DB 10;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275
```

```
RESULT 14
US-10-352-684A-48
; Sequence 48, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-352-684A-48

  Query Match             98.2%;  Score 1368;  DB 15;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275
```

```
RESULT 15
US-10-174-364-44
; Sequence 44, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson et al.
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-729CIP2
;  CURRENT APPLICATION NUMBER: US/10/174,364
;  CURRENT FILING DATE:  2002-06-17
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-174-364-44

  Query Match             98.2%;  Score 1368;  DB 15;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-122;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275

Search completed: August 27, 2005, 23:24:28
Job time : 115 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         August 27, 2005, 21:22:59 ; Search time 62 Seconds
                                          (without alignments)
                                          299.800 Million cell updates/sec

Title:          US-09-598-982C-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:      513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                 %
Result           Query
   No.    Score  Match Length DB  ID                          Description
 ----------------------------------------------------------------------------
     1     1387   99.6    249  3  US-09-079-970A-5             Sequence 5, Appli
     2     1368   98.2    245  3  US-09-079-970A-6             Sequence 6, Appli
     3     1368   98.2    245  4  US-09-601-318-1              Sequence 1, Appli
     4     1368   98.2    274  2  US-09-016-366A-21            Sequence 21, Appl
     5     1368   98.2    274  2  US-08-978-404B-16            Sequence 16, Appl
     6     1363   97.8    273  2  US-09-016-366A-19            Sequence 19, Appl
     7     1363   97.8    273  2  US-08-978-404B-14            Sequence 14, Appl
     8     1361   97.7    244  4  US-09-601-318-4              Sequence 4, Appli
     9     1361   97.7    244  4  US-09-601-318-5              Sequence 5, Appli
    10     1361   97.7    244  4  US-09-601-318-6              Sequence 6, Appli
    11     1361   97.7    244  4  US-09-601-318-7              Sequence 7, Appli
    12     1344   96.5    267  2  US-09-016-366A-23            Sequence 23, Appl
    13     1344   96.5    267  2  US-08-978-404B-18            Sequence 18, Appl
    14     1344   96.5    267  4  US-09-917-254-101            Sequence 101, App
    15     1274   91.5    245  3  US-08-944-483-69             Sequence 69, Appl
    16     1274   91.5    275  2  US-09-016-366A-17            Sequence 17, Appl
    17     1274   91.5    275  2  US-08-978-404B-12            Sequence 12, Appl
    18     1091   78.3    276  2  US-09-016-366A-15            Sequence 15, Appl
    19     1091   78.3    276  2  US-08-978-404B-21            Sequence 21, Appl
    20     1075   77.2    270  2  US-08-978-404B-8             Sequence 8, Appli
    21     1054   75.7    273  2  US-08-978-404B-3             Sequence 3, Appli
    22     1051   75.4    274  2  US-08-978-404B-5             Sequence 5, Appli
    23     1037   74.4    273  2  US-08-978-404B-6             Sequence 6, Appli
    24      702   50.4    190  2  US-08-845-998-4              Sequence 4, Appli
    25      702   50.4    190  3  US-09-206-537-4              Sequence 4, Appli
    26      702   50.4    190  3  US-09-430-854-4              Sequence 4, Appli
    27      698   50.1    190  2  US-08-845-998-6              Sequence 6, Appli
```

```
    28      698   50.1    190  3  US-09-206-537-6              Sequence 6, Appli
    29      698   50.1    190  3  US-09-430-854-6              Sequence 6, Appli
    30    684.5   49.1    269  2  US-08-978-404B-10            Sequence 10, Appl
    31    563.5   40.5    290  4  US-09-386-653A-7             Sequence 7, Appli
    32    554.5   39.8    315  4  US-09-386-653A-9             Sequence 9, Appli
    33    536.5   38.5    284  4  US-09-387-375-7              Sequence 7, Appli
    34    536.5   38.5    284  4  US-10-041-400A-7             Sequence 7, Appli
    35    536.5   38.5    284  4  US-10-042-091A-7             Sequence 7, Appli
    36    527.5   37.9    316  4  US-09-387-375-9              Sequence 9, Appli
    37    527.5   37.9    316  4  US-10-041-400A-9             Sequence 9, Appli
    38    527.5   37.9    316  4  US-10-042-091A-9             Sequence 9, Appli
    39      507   36.4    319  4  US-09-386-642-12             Sequence 12, Appl
    40      507   36.4    328  4  US-09-386-642-11             Sequence 11, Appl
    41      506   36.3    299  3  US-08-944-483-66             Sequence 66, Appl
    42    501.5   36.0    317  4  US-09-386-629-7              Sequence 7, Appli
    43    501.5   36.0    317  4  US-09-907-794A-263           Sequence 263, App
    44    501.5   36.0    317  4  US-09-905-125A-263           Sequence 263, App
    45    501.5   36.0    317  4  US-09-902-775A-263           Sequence 263, App

                                  ALIGNMENTS


RESULT 1
US-09-079-970A-5
; Sequence 5, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  249 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-079-970A-5

  Query Match             99.6%;  Score 1387;  DB 3;  Length 249;
  Best Local Similarity   99.6%;  Pred. No. 7.8e-141;
  Matches  248;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
                ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60
```

```
Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 2
US-09-079-970A-6
; Sequence 6, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  245 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-079-970A-6

  Query Match             98.2%;  Score 1368;  DB 3;  Length 245;
  Best Local Similarity   99.6%;  Pred. No. 8.4e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245


RESULT 3
US-09-601-318-1
; Sequence 1, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 245
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-1

  Query Match             98.2%;  Score 1368;  DB 4;  Length 245;
  Best Local Similarity   99.6%;  Pred. No. 8.4e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245
```

```
RESULT 4
US-09-016-366A-21
; Sequence 21, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-21

  Query Match             98.2%;  Score 1368;  DB 2;  Length 274;
  Best Local Similarity   99.6%;  Pred. No. 9.8e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         30 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 269

Qy        245 VPKKP 249
              |||||
Db        270 VPKKP 274


RESULT 5
US-08-978-404B-16
; Sequence 16, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-16

  Query Match             98.2%;  Score 1368;  DB 2;  Length 274;
  Best Local Similarity   99.6%;  Pred. No. 9.8e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         30 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 269

Qy        245 VPKKP 249
              |||||
Db        270 VPKKP 274


RESULT 6
US-09-016-366A-19
; Sequence 19, Application US/09016366A
; Patent No. 5955431
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-19

  Query Match             97.8%;  Score 1363;  DB 2;  Length 273;
  Best Local Similarity   99.2%;  Pred. No. 3.4e-138;
  Matches  243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         29 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         89 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 268

Qy        245 VPKKP 249
              |||||
Db        269 VPKKP 273


RESULT 7
US-08-978-404B-14
; Sequence 14, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-14

  Query Match             97.8%;  Score 1363;  DB 2;  Length 273;
  Best Local Similarity   99.2%;  Pred. No. 3.4e-138;
  Matches  243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         29 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         89 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 268

Qy        245 VPKKP 249
              |||||
Db        269 VPKKP 273


RESULT 8
US-09-601-318-4
; Sequence 4, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
```

```
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-4

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 9
US-09-601-318-5
; Sequence 5, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
```

```
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-5

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 10
US-09-601-318-6
; Sequence 6, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
```

```
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-6

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 11
US-09-601-318-7
; Sequence 7, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
US-09-601-318-7

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 4.7e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 12
US-09-016-366A-23
; Sequence 23, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-23

  Query Match             96.5%;  Score 1344;  DB 2;  Length 267;
```

```
  Best Local Similarity   98.4%;  Pred. No. 3.6e-136;
  Matches  241;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db         23 IVGGQEAPRSKWPWQVSLRVDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 82

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 142

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 202

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 262

Qy        245 VPKKP 249
              |||||
Db        263 VPKKP 267


RESULT 13
US-08-978-404B-18
; Sequence 18, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-18

  Query Match             96.5%;  Score 1344;  DB 2;  Length 267;
  Best Local Similarity   98.4%;  Pred. No. 3.6e-136;
  Matches  241;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db         23 IVGGQEAPRSKWPWQVSLRVDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 82

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 142

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 202

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 262

Qy        245 VPKKP 249
              |||||
Db        263 VPKKP 267


RESULT 14
US-09-917-254-101
; Sequence 101, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Baak, Jan
;  TITLE OF INVENTION: Prognostic Classification of Breast Cancer
;  FILE REFERENCE: B0801/7224(JRV)
;  CURRENT APPLICATION NUMBER: US/09/917,254
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/222,093
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 101
;   LENGTH: 267
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-917-254-101

  Query Match             96.5%;  Score 1344;  DB 4;  Length 267;
  Best Local Similarity   98.4%;  Pred. No. 3.6e-136;
  Matches  241;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db         23 IVGGQEAPRSKWPWQVSLRVDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 82

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 142

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 202

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 262

Qy        245 VPKKP 249
              |||||
Db        263 VPKKP 267


RESULT 15
US-08-944-483-69
; Sequence 69, Application US/08944483
; Patent No. 6232456
;  GENERAL INFORMATION:
;    APPLICANT:  COHEN, MAURICE
```

```
;    APPLICANT:  COLPITTS, TRACEY L.
;    APPLICANT:  FRIEDMAN, PAULA N.
;    APPLICANT:  GRANADOS, EDWARD N.
;    APPLICANT:  KLASS, MICHAEL R.
;    APPLICANT:  RUSSELL, JOHN C.
;    APPLICANT:  STEWART, KENT D.
;    APPLICANT:  STROUPE, STEVEN D.
;    TITLE OF INVENTION:  NOVEL SERINE PROTEASE REAGENTS
;    TITLE OF INVENTION:  AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
;    TITLE OF INVENTION:  OF THE PROSTATE
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,483
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6183.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  245 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6232456e
US-08-944-483-69

  Query Match             91.5%;  Score 1274;  DB 3;  Length 245;
  Best Local Similarity   92.2%;  Pred. No. 1.1e-128;
  Matches  226;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||||
Db          1 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        121 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245


Search completed: August 27, 2005, 23:13:02
```

```
Job time : 64 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:         August 27, 2005, 19:27:19 ; Search time 626 Seconds
                                             (without alignments)
                                            8058.662 Million cell updates/sec

Title:          US-09-598-982C-20
Perfect score:  771
Sequence:       1 gggcccctcgagaaaagaat.........cgtgaagcggccgccgtcgt 771

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters:      14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
                19:  /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
                20:  /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
                21:  /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
                22:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                23:  /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
                24:  /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
                25:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                26:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                    %
Result              Query
  No.     Score     Match Length DB  ID                          Description
 ---------------------------------------------------------------------------
     1      731     94.8    1143  17  US-10-352-684A-47            Sequence 47, Appl
     2      731     94.8    1145  18  US-10-287-226-93             Sequence 93, Appl
     3      731     94.8    1148  21  US-10-956-157-2444           Sequence 2444, Ap
     4      731     94.8    1193  21  US-10-956-157-1043           Sequence 1043, Ap
     5    730.4     94.7     828  18  US-10-287-226-91             Sequence 91, Appl
     6    729.4     94.6    1334  20  US-10-723-860-6799           Sequence 6799, Ap
     7    726.2     94.2    1081  9  US-09-954-456-2126           Sequence 2126, Ap
     8    726.2     94.2    1081  10  US-09-960-706-680            Sequence 680, App
     9    726.2     94.2    1081  10  US-09-873-319-427            Sequence 427, App
    10    726.2     94.2    1081  10  US-09-873-367C-155           Sequence 155, App
    11    726.2     94.2    1081  10  US-09-873-367C-714           Sequence 714, App
    12    726.2     94.2    1081  21  US-10-843-641A-155           Sequence 155, App
    13    726.2     94.2    1081  21  US-10-843-641A-714           Sequence 714, App
    14    726.2     94.2    1081  21  US-10-843-641A-5153          Sequence 5153, Ap
    15    718.2     93.2    2662  18  US-10-275-505-27             Sequence 27, Appl
    16      700     90.8     828  18  US-10-287-226-89             Sequence 89, Appl
    17    686.2     89.0    1154  17  US-10-352-684A-45            Sequence 45, Appl
    18    686.2     89.0    1158  14  US-10-116-802-240            Sequence 240, App
    19    686.2     89.0    1158  21  US-10-956-157-1042           Sequence 1042, Ap
    20      678     87.9     858  18  US-10-287-226-87             Sequence 87, Appl
    21    507.6     65.8     729  21  US-10-480-988-55             Sequence 55, Appl
    22      498     64.6     729  21  US-10-956-157-1789           Sequence 1789, Ap
    23    484.4     62.8     600  21  US-10-956-157-7024           Sequence 7024, Ap
    24      463     60.1    1187  18  US-10-275-505-22             Sequence 22, Appl
    25      380     49.3    5456  20  US-10-723-860-2689           Sequence 2689, Ap
    26      380     49.3    5456  22  US-10-756-149-2484           Sequence 2484, Ap
    27    287.6     37.3     600  21  US-10-956-157-6278           Sequence 6278, Ap
    28    287.6     37.3     600  21  US-10-956-157-7679           Sequence 7679, Ap
    29    287.4     37.3     846  10  US-09-813-432-11             Sequence 11, Appl
    30    287.4     37.3     846  17  US-10-174-364-11             Sequence 11, Appl
    31    287.4     37.3     846  18  US-10-246-583-11             Sequence 11, Appl
    32    287.4     37.3     846  19  US-10-689-832-11             Sequence 11, Appl
    33      275     35.7     383  10  US-09-918-995-6378           Sequence 6378, Ap
    34    274.8     35.6     600  21  US-10-956-157-6277           Sequence 6277, Ap
    35    252.4     32.7     948  16  US-10-131-409-15             Sequence 15, Appl
    36    252.4     32.7     948  17  US-10-139-854-15             Sequence 15, Appl
    37    252.4     32.7     948  17  US-10-150-813-15             Sequence 15, Appl
    38    252.4     32.7     948  17  US-10-150-811-15             Sequence 15, Appl
    39    236.4     30.7     858  16  US-10-131-409-21             Sequence 21, Appl
    40    236.4     30.7     858  17  US-10-139-854-21             Sequence 21, Appl
    41    236.4     30.7     858  17  US-10-150-813-21             Sequence 21, Appl
    42    236.4     30.7     858  17  US-10-150-811-21             Sequence 21, Appl
    43    233.4     30.3     873  14  US-10-117-323-2              Sequence 2, Appli
    44    231.6     30.0     843  16  US-10-131-409-25             Sequence 25, Appl
    45    231.6     30.0     843  17  US-10-139-854-25             Sequence 25, Appl


                                   ALIGNMENTS


RESULT 1
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
;   LENGTH: 1143
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6)...(833)
US-10-352-684A-47

  Query Match             94.8%;  Score 731;  DB 17;  Length 1143;
  Best Local Similarity   98.0%;  Pred. No. 2.6e-195;
  Matches  740;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 153

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 213

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        214 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 273

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 333

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 393

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 453

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 513

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 573

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 633

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 693

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 753

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 813

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              ||||||||||||||||||||  | | |   |  ||
Db        814 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 848

RESULT 2
US-10-287-226-93
; Sequence 93, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
```

```
; SEQ ID NO 93
;   LENGTH: 1145
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(832)
US-10-287-226-93

  Query Match             94.8%;  Score 731;  DB 18;  Length 1145;
  Best Local Similarity   98.7%;  Pred. No. 2.6e-195;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           96 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 155

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          156 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 215

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db          216 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 275

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          276 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 335

Qy          257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          336 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 395

Qy          317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          396 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 455

Qy          377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          456 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 515

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          516 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 575

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          576 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 635

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          636 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 695

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          696 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 755

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          756 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 815

Qy          737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
                ||||||||||||||||||||  | | |
Db          816 ATGTCCCCAAAAAGCCGTGAGTCAGGC 842

RESULT 3
US-10-956-157-2444
; Sequence 2444, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
```

```
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2444
;   LENGTH: 1148
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-2444

  Query Match             94.8%;  Score 731;  DB 21;  Length 1148;
  Best Local Similarity   98.7%;  Pred. No. 2.6e-195;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           96 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 155

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          156 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 215

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db          216 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 275

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          276 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 335

Qy          257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          336 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 395

Qy          317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          396 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 455

Qy          377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          456 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 515

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          516 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 575

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          576 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 635

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          636 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 695

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          696 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 755

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          756 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 815

Qy          737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
                ||||||||||||||||||||  | | |
Db          816 ATGTCCCCAAAAAGCCGTGAGTCAGGC 842

RESULT 4
US-10-956-157-1043
; Sequence 1043, Application US/10956157
```

```
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1043
;   LENGTH: 1193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-1043

  Query Match             94.8%;  Score 731;  DB 21;  Length 1193;
  Best Local Similarity   98.7%;  Pred. No. 2.6e-195;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         141 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 200

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 260

Qy         137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db         261 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 320

Qy         197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         321 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 380

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         381 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 440

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 500

Qy         377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         501 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 560

Qy         437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         561 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 620

Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         621 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 680

Qy         557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         681 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 740

Qy         617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         741 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 800

Qy         677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         801 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 860

Qy         737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
               ||||||||||||||||||||  | | |
Db         861 ATGTCCCCAAAAAGCCGTGAGTCAGGC 887
```

```
RESULT 5
US-10-287-226-91
; Sequence 91, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 91
;   LENGTH: 828
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-91

  Query Match             94.7%;  Score 730.4;  DB 18;  Length 828;
  Best Local Similarity   99.2%;  Pred. No. 3.7e-195;
  Matches  734;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 148

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 208

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        209 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 268

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 328

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 388

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 448

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 508

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 568

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 628

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 688

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 748

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 808

Qy        737 ATGTCCCCAAAAAGCCGTGA 756
              ||||||||||||||||||||
Db        809 ATGTCCCCAAAAAGCCGTGA 828

RESULT 6
US-10-723-860-6799
; Sequence 6799, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6799
;   LENGTH: 1334
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-6799

  Query Match             94.6%;  Score 729.4;  DB 20;  Length 1334;
  Best Local Similarity   98.5%;  Pred. No. 7.5e-195;
  Matches  736;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 200

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TCCGCGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 260

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 320

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 380

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 440

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 500

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 560

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 620

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 680

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 740

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 800

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 860

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        861 ATGTCCCCAAAAAGCCGTGAGTCAGGC 887

RESULT 7
US-09-954-456-2126
; Sequence 2126, Application US/09954456
```

```
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-2126

  Query Match             94.2%;  Score 726.2;  DB 9;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 8
US-09-960-706-680
; Sequence 680, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasi
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
US-09-960-706-680

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425
```

```
Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 9
US-09-873-319-427
; Sequence 427, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 427
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
US-09-873-319-427

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 10
US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-155

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
```

```
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 11
US-09-873-367C-714
; Sequence 714, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
```

```
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 714
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-714

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 12
```

```
US-10-843-641A-155
; Sequence 155, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-155

  Query Match             94.2%;  Score 726.2;  DB 21;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 13
US-10-843-641A-714
; Sequence 714, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 714
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-714

  Query Match             94.2%;  Score 726.2;  DB 21;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 14
US-10-843-641A-5153
; Sequence 5153, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5153
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-5153

  Query Match             94.2%;  Score 726.2;  DB 21;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 5.8e-194;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 15
US-10-275-505-27
```

```
; Sequence 27, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 27
;   LENGTH: 2662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7472460CB1
US-10-275-505-27

  Query Match             93.2%;  Score 718.2;  DB 18;  Length 2662;
  Best Local Similarity   98.2%;  Pred. No. 1.1e-191;
  Matches  726;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 758

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 818

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        819 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 878

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 938

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 998

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 1058

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 1118

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       1119 CACCGCCATTTCCTCTGAAGCATGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 1178

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 1238

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 1298

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||
Db       1299 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGATGGGGAGAGGGCTGTGCCC 1358

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1359 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 1418

Qy        737 ATGTCCCCAAAAAGCCGTG 755
              |||||||||||||||||||
Db       1419 ATGTCCCCAAAAAGCCGTG 1437

Search completed: August 27, 2005, 21:22:53
Job time : 630 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         August 27, 2005, 18:58:19 ; Search time 188 Seconds
                                          (without alignments)
                                          6710.475 Million cell updates/sec

Title:          US-09-598-982C-20
Perfect score:  771
Sequence:       1 gggcccctcgagaaaagaat..........cgtgaagcggccgccgtcgt 771

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:      2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                        Description
---------------------------------------------------------------------------
    1   764.6    99.2    771  3  US-09-079-970A-4           Sequence 4, Appli
    2     731    94.8   1128  2  US-09-016-366A-20          Sequence 20, Appl
    3     731    94.8   1128  2  US-08-978-404B-15          Sequence 15, Appl
    4     731    94.8   1137  2  US-09-016-366A-18          Sequence 18, Appl
    5     731    94.8   1137  2  US-08-978-404B-13          Sequence 13, Appl
    6   728.6    94.5    735  3  US-09-079-970A-1           Sequence 1, Appli
    7   726.2    94.2   1081  2  US-09-016-366A-22          Sequence 22, Appl
    8   726.2    94.2   1081  2  US-08-978-404B-17          Sequence 17, Appl
    9   726.2    94.2   1081  4  US-09-917-254-50           Sequence 50, Appl
   10   686.2    89.0   1154  2  US-09-016-366A-16          Sequence 16, Appl
   11   686.2    89.0   1154  2  US-08-978-404B-11          Sequence 11, Appl
   12   477.2    61.9   1219  2  US-08-978-404B-7           Sequence 7, Appli
   13   469.6    60.9   1108  2  US-09-016-366A-14          Sequence 14, Appl
   14   469.6    60.9   1108  2  US-08-978-404B-20          Sequence 20, Appl
   15     463    60.1   1031  2  US-08-978-404B-1           Sequence 1, Appli
   16   445.6    57.8   1103  2  US-09-016-366A-24          Sequence 24, Appl
   17     444    57.6   1097  2  US-08-978-404B-4           Sequence 4, Appli
   18     371    48.1   2259  2  US-08-845-998-3            Sequence 3, Appli
   19     371    48.1   2259  3  US-09-206-537-3            Sequence 3, Appli
   20     371    48.1   2259  3  US-09-430-854-3            Sequence 3, Appli
   21   369.4    47.9   2218  2  US-08-845-998-5            Sequence 5, Appli
   22   369.4    47.9   2218  3  US-09-206-537-5            Sequence 5, Appli
   23   369.4    47.9   2218  3  US-09-430-854-5            Sequence 5, Appli
   24   271.8    35.3   1095  2  US-08-978-404B-9           Sequence 9, Appli
   25   201.4    26.1    980  4  US-09-023-942A-30          Sequence 30, Appl
   26   201.4    26.1   1110  3  US-09-386-653A-1           Sequence 1, Appli
   27   199.8    25.9   1212  4  US-09-620-312D-431         Sequence 431, App
   28   193.8    25.1   1130  3  US-09-386-653A-8           Sequence 8, Appli
   29   187.8    24.4   1613  4  US-09-387-375-1            Sequence 1, Appli
   30   187.8    24.4   1613  4  US-10-041-400A-1           Sequence 1, Appli
   31   187.8    24.4   1613  4  US-10-042-091A-1           Sequence 1, Appli
   32   181.2    23.5   1130  4  US-09-387-375-8            Sequence 8, Appli
   33   181.2    23.5   1130  4  US-10-041-400A-8           Sequence 8, Appli
   34   181.2    23.5   1130  4  US-10-042-091A-8           Sequence 8, Appli
   35     153    19.8   3757  2  US-09-016-366A-13          Sequence 13, Appl
   36     153    19.8   3757  2  US-08-978-404B-19          Sequence 19, Appl
   37     152    19.7    933  4  US-09-023-942A-29          Sequence 29, Appl
   38   143.8    18.7   2397  2  US-08-978-404B-2           Sequence 2, Appli
   39   140.6    18.2   1142  3  US-09-386-642-8            Sequence 8, Appli
   40   140.6    18.2   1169  3  US-09-386-642-7            Sequence 7, Appli
   41   132.2    17.1   1154  4  US-09-636-382A-1           Sequence 1, Appli
   42   131.2    17.0   1165  4  US-09-023-942A-28          Sequence 28, Appl
   43   131.2    17.0   1378  4  US-09-907-794A-262         Sequence 262, App
   44   131.2    17.0   1378  4  US-09-905-125A-262         Sequence 262, App
   45   131.2    17.0   1378  4  US-09-902-775A-262         Sequence 262, App

                               ALIGNMENTS

RESULT 1
US-09-079-970A-4
; Sequence 4, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  7..753
;    FEATURE:
;      NAME/KEY:  misc_signal
```

```
;      LOCATION:  7..18
US-09-079-970A-4

  Query Match             99.2%;  Score 764.6;  DB 3;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 1.4e-169;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 2
US-09-016-366A-20
; Sequence 20, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
```

```
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-20

  Query Match             94.8%;  Score 731;  DB 2;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-161;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626
```

```
Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833

RESULT 3
US-08-978-404B-15
; Sequence 15, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-15

  Query Match             94.8%;  Score 731;  DB 2;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-161;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206
```

```
Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833

RESULT 4
US-09-016-366A-18
; Sequence 18, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
```

```
;     FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-18

  Query Match             94.8%;  Score 731;  DB 2;  Length 1137;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-161;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 144

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 204

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 264

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 324

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 384

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 444

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 504

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 564

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 624

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 684

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 744

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 804

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        805 ATGTCCCCAAAAAGCCGTGAGTCAGGC 831
```

```
RESULT 5
US-08-978-404B-13
; Sequence 13, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-13

  Query Match             94.8%;  Score 731;  DB 2;  Length 1137;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-161;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 144

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 204

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 264

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 324

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 384

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 444
```

```
Qy          377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          445 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 504

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          505 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 564

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          565 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 624

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          625 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 684

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          685 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 744

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          745 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 804

Qy          737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
                ||||||||||||||||||||  | | |
Db          805 ATGTCCCCAAAAAGCCGTGAGTCAGGC 831

RESULT 6
US-09-079-970A-1
; Sequence 1, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
```

```
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..735
US-09-079-970A-1

  Query Match             94.5%;  Score 728.6;  DB 3;  Length 735;
  Best Local Similarity   99.5%;  Pred. No. 3.7e-161;
  Matches  731;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy           19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy           79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy          139 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198
                ||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy          199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy          259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy          319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy          379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy          439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy          499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy          559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy          619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy          679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy          739 GTCCCCAAAAAGCCG 753
                |||||||||||||||
Db          721 GTCCCCAAAAAGCCG 735

RESULT 7
US-09-016-366A-22
; Sequence 22, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
```

```
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-22

  Query Match             94.2%;  Score 726.2;  DB 2;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 1.5e-160;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545
```

```
Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 8
US-08-978-404B-17
; Sequence 17, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-17

  Query Match             94.2%;  Score 726.2;  DB 2;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 1.5e-160;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125
```

```
Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 9
US-09-917-254-50
; Sequence 50, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Baak, Jan
;  TITLE OF INVENTION: Prognostic Classification of Breast Cancer
;  FILE REFERENCE: B0801/7224(JRV)
;  CURRENT APPLICATION NUMBER: US/09/917,254
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/222,093
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-917-254-50

  Query Match             94.2%;  Score 726.2;  DB 4;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 1.5e-160;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;
```

```
Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 10
US-09-016-366A-16
; Sequence 16, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
```

```
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-16

  Query Match             89.0%;  Score 686.2;  DB 2;  Length 1154;
  Best Local Similarity   94.3%;  Pred. No. 3.3e-151;
  Matches  712;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          106 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 165

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          166 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 225

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                ||||||| |||   ||| |||||||||||||||||||||||||| |||||||||||||||
Db          226 TGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 285

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          286 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGC 345

Qy          257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
                ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db          346 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 405

Qy          317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
                | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db          406 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 465

Qy          377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
                |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          466 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 525

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          526 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 585

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db          586 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 645

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||||||||||| || ||||||||| ||| |||||||||| ||||||||||||||||
Db          646 TGTGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGT 705

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                |||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db          706 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCC 765

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          766 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 825

Qy          737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                ||||||||||||||||||||  | | |   |  ||
Db          826 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 860


RESULT 11
US-08-978-404B-11
; Sequence 11, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-11

  Query Match             89.0%;  Score 686.2;  DB 2;  Length 1154;
  Best Local Similarity   94.3%;  Pred. No. 3.3e-151;
  Matches  712;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          106 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 165

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          166 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 225

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                ||||||| |||   ||| |||||||||||||||||||||||||| |||||||||||||||
Db          226 TGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 285

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
```

```
Db        286 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGC 345

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db        346 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 405

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db        406 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 465

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        466 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 525

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 585

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        586 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 645

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||| || ||||||||| ||| |||||||||| ||||||||||||||||
Db        646 TGTGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGT 705

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db        706 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCC 765

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 825

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              ||||||||||||||||||||  | | |   |  ||
Db        826 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 860

RESULT 12
US-08-978-404B-7
; Sequence 7, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1219 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-7

  Query Match             61.9%;  Score 477.2;  DB 2;  Length 1219;
  Best Local Similarity   77.1%;  Pred. No. 2.5e-102;
  Matches  581;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

Qy          6 CCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGT 65
              ||  |||    | || || ||||| |||||||| ||  ||| ||||||||||||||||||
Db        260 CCAAGAGTGGGGCATTGTTGGGGGACAGGAGGCACCTGGGAACAAGTGGCCCTGGCAGGT 319

Qy         66 GAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCC 125
              ||||||  | | | | |   | ||||||| ||| |||||||| |||||||||||||||||
Db        320 GAGCCTTCGTGCCAATGAAACCTACTGGAGGCATTTCTGCGGCGGCTCCCTCATCCACCC 379

Qy        126 CCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCT 185
               ||||||||||| ||||| ||    || |||||||||    |    ||||    |    |
Db        380 ACAGTGGGTGCTCACCGCGGCACACTGTGTGGGACCGACTATTGCTGATCCCAACAAGGT 439

Qy        186 CAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAG 245
              ||| || || || ||  ||||| ||||||| ||||| ||||| |||||| | || ||| |
Db        440 CAGAGTACAGCTTCGAAAGCAGTACCTCTATTACCACGACCACCTGCTGGCTGTGAGCCG 499

Qy        246 GATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA 305
              |||||||   |||||    |||||  || ||||||  || |||||||||||||| || ||
Db        500 GATCATCACACACCCGACATTCTATGCCACCCAGAATGGGGCGGACATCGCCCTACTTGA 559

Qy        306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTC 365
              |||  || | || || ||  | |||||||| |||||| | ||| ||||||| ||||||||
Db        560 GCTCAAGAACCCTGTAAACATTTCCAGCCATGTCCACCCCGTCTCCCTGCCTCCTGCCTC 619

Qy        366 AGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGA 425
              ||||||||||||| | || |    ||||||||| || ||||||||  |  | ||||||||
Db        620 AGAGACCTTCCCCTCAGGAACATTGTGCTGGGTGACAGGCTGGGGAAACATCGACAATGA 679

Qy        426 TGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCA 485
              || | |||| ||||||||||||||  ||||| ||||| | || ||| |  ||||||||||
Db        680 TGTGAGCCTGCCACCGCCATTTCCCTTGAAGGAGGTGCAAGTTCCCGTCGTGGAAAACCA 739

Qy        486 CATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCG 545
                ||||||||   || || |||   || | |||||| || ||| || ||| ||| |||||
Db        740 GCTTTGTGACCTGAAGTATCACAAAGGTGTCTACACAGGGGACAACATCCACATTGTCCG 799

Qy        546 TGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGG 605
               ||||||||||||||||| ||||||    |    ||||| |||||||| |||||||||||
Db        800 AGACGACATGCTGTGTGCTGGGAACGAAGGACACGACTCCTGCCAGGGTGACTCCGGAGG 859

Qy        606 GCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGA 665
               || ||||| |||||||| || || |||||||||||||| || |||||||||||||| ||
Db        860 ACCTCTGGTCTGCAAGGTAAACGGTACCTGGCTGCAGGCAGGTGTGGTCAGCTGGGGTGA 919

Qy        666 GGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTG 725
              ||||||||| | ||||||| |||||||||||||||| || |||||||| |||||||||||
Db        920 GGGCTGTGCTCTGCCCAACAGGCCTGGCATCTACACTCGGGTCACCTATTACTTGGACTG 979

Qy        726 GATCCACCACTATGTCCCCAAAAAGCCGTGAAGC 759
              |||||||| ||||||||||||  |    |||| |
Db        980 GATCCACCGCTATGTCCCCAAGGACTTCTGAATC 1013

RESULT 13
US-09-016-366A-14
; Sequence 14, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
```

```
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-14

  Query Match             60.9%;  Score 469.6;  DB 2;  Length 1108;
  Best Local Similarity   77.2%;  Pred. No. 1.5e-100;
  Matches  571;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| || || || |||||  |   ||| |||||||||||||||||||||||||||
Db        125 GCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 184

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              |  |       |||||||| || |||||||| ||||| ||||||||||| ||||||||||
Db        185 TTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGC 244

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              | || || ||    || ||||||||| || ||||    |       | || ||||||| |
Db        245 TCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGC 304

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              | || |||||| | || |||||   ||||||||| ||  |  | | | ||||| | ||||
Db        305 TTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGC 364

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              |||| || |  ||||| |||  ||   || || ||| | |||||||||||||| ||||  |
Db        365 ACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCC 424

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | ||||| ||||||| |||  ||||| |  |  |||||||||||||||| ||||||||||
Db        425 CTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCC 484

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||| ||||   |||||||||| || |||||||||||  |  || |||||  ||||  ||||
Db        485 CCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCC 544

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              |||| || | |||||||||||| |||||||| |||||  ||||||||  | | ||||||
Db        545 CACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACC 604

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                || ||||||  ||||  |||||||||||| ||  | | ||| |||| ||| | |||||
Db        605 GGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGC 664

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||| || || ||| |||| ||||| ||||||||||| || || ||||| ||||| |
Db        665 TGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCT 724

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||| ||||| || |||||||||||||| || |||||||||||||| |||||||| || |
Db        725 GCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCAC 784

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||  |||||||||||||||||||| || || |||||||| |||||||||||||| ||
Db        785 AGCCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCT 844

Qy        737 ATGTCCCCAAAAAGCCGTGA 756
              |||||||  |  |  | |||
Db        845 ATGTCCCTGAGCATTCCTGA 864


RESULT 14
US-08-978-404B-20
; Sequence 20, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-20

  Query Match             60.9%;  Score 469.6;  DB 2;  Length 1108;
  Best Local Similarity   77.2%;  Pred. No. 1.5e-100;
  Matches  571;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
```

```
              | ||||| || || || |||||  |   ||| |||||||||||||||||||||||||||
Db        125 GCATCGTGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 184

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              |  |       |||||||| || |||||||| ||||| ||||||||||| ||||||||||
Db        185 TTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGC 244

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              | || || ||    || ||||||||| || ||||    |       | || ||||||| |
Db        245 TCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGC 304

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              | || |||||| | || |||||   ||||||||| ||  |  | | | ||||| | ||||
Db        305 TTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGC 364

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              |||| || |  ||||| ||| ||   || || ||| | |||||||||||||| ||||  |
Db        365 ACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCC 424

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | ||||| ||||||| |||  ||||| |  |  |||||||||||||||| ||||||||||
Db        425 CTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCC 484

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||| ||||   |||||||||| || |||||||||||  | || ||||| ||||  ||||
Db        485 CCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCC 544

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              |||| || | |||||||||||| |||||||| |||||  ||||||||  | | ||||||
Db        545 CACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACC 604

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                || ||||||  ||||  |||||||||||| ||  | | ||| |||| ||| | |||||
Db        605 GGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGC 664

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||| || || ||| |||| ||||| ||||||||||| || || ||||| ||||| |
Db        665 TGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCT 724

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||| ||||| || |||||||||||||| || |||||||||||||| |||||||| || |
Db        725 GCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCAC 784

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||  |||||||||||||||||| || || |||||||| ||||||||||||| ||
Db        785 AGCCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCT 844

Qy        737 ATGTCCCCAAAAAGCCGTGA 756
              |||||||  |  |  | |||
Db        845 ATGTCCCTGAGCATTCCTGA 864

RESULT 15
US-08-978-404B-1
; Sequence 1, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-1

  Query Match             60.1%;  Score 463;  DB 2;  Length 1031;
  Best Local Similarity   76.4%;  Pred. No. 5.1e-99;
  Matches  568;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy         14 AAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGA 73
              || | || || ||||| |||||||| |   ||| |||||||||||||||||||||||||
Db        106 AAGGCATTGTGGGGGGACAGGAGGCACATGGGAACAAGTGGCCCTGGCAGGTGAGCCTGC 165

Qy         74 GAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGG 133
              | | | | | | | ||||||||||| |||||||| ||||||||||||||||| |||||||
Db        166 GTGCCAATGACACCTACTGGATGCATTTCTGCGGTGGCTCCCTCATCCACCCACAGTGGG 225

Qy        134 TGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGC 193
              |||| || || ||    || ||||||||||| ||    || |    |    |||| || |
Db        226 TGCTCACTGCGGCACACTGTGTGGGACCGGATGTTGCTGACCCCAACAAGGTCAGAGTAC 285

Qy        194 AACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCG 253
              | || ||  ||||| ||||||| ||||| ||||| ||| || | || |||  |||||||
Db        286 AGCTCCGTAAGCAGTACCTCTATTACCATGACCACCTGATGACTGTGAGCCAGATCATCA 345

Qy        254 TGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGG 313
                |||||  | ||||||| || ||||   || || ||||| ||||||||| | ||
Db        346 CACACCCCGACTTCTACATCGTCCAGGATGGGGCAGACATTGCCCTGCTGAAACTCACAA 405

Qy        314 AGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCT 373
              | || |||||  | ||   | | |||||| | ||| |||| || ||||||||||||||||
Db        406 ACCCTGTGAACATTTCTGACTATGTCCACCCTGTCCCCCTACCTCCTGCCTCAGAGACCT 465

Qy        374 TCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCC 433
              ||||| | || | |  ||||||||| || ||||||||  |  | ||||||| ||    ||
Db        466 TCCCCTCAGGAACGTTGTGCTGGGTGACAGGCTGGGGTAACATCGACAATGGTGTAAACC 525

Qy        434 TCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTG 493
              | || || ||||||||| ||||| ||||| | || ||||| || ||||||||| ||||||
Db        526 TGCCGCCACCATTTCCTTTGAAGGAGGTGCAAGTTCCCATTATAGAAAACCACCTTTGTG 585

Qy        494 ACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACA 553
              ||   || || |||   ||   |  ||| || ||| | |||| ||| ||||| || ||||
Db        586 ACTTGAAGTATCACAAAGGTCTCATCACAGGTGACAATGTCCACATTGTCCGAGATGACA 645

Qy        554 TGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGG 613
              |||||||||| |||||     |    ||||| |||||||||||||||||||| || ||||
Db        646 TGCTGTGTGCTGGGAATGAAGGACATGACTCCTGCCAGGGCGACTCCGGAGGACCTCTGG 705

Qy        614 TGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTG 673
              | ||||||||  | | ||||||||||||||| ||||||||||||||||| ||||||||||
Db        706 TCTGCAAGGTAGAAGACACCTGGCTGCAGGCAGGCGTGGTCAGCTGGGGTGAGGGCTGTG 765

Qy        674 CCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACC 733
              | ||||||||| ||||||||||||||||||| |||||||| |||||||||||||||||||
```

```
Db        766 CACAGCCCAACAGGCCTGGCATCTACACCCGGGTCACCTATTACTTGGACTGGATCCACC 825

Qy        734 ACTATGTCCCCAAAAAGCCGTGA 756
              |||||||||||||  |    |||
Db        826 ACTATGTCCCCAAGGACTTCTGA 848

Search completed: August 27, 2005, 19:37:45
Job time : 191 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        August 27, 2005, 23:19:35 ; Search time 56 Seconds
                                           (without alignments)
                                           472.493 Million cell updates/sec

Title:         US-09-598-982C-52
Perfect score: 1512
Sequence:      1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:      283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_79:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                     %
Result            Query
  No.   Score     Match Length DB ID                  Description
--------------------------------------------------------------------------
    1  1470.5    97.3    274  2  A45754              tryptase (EC 3.4.2
    2    1430    94.6    275  2  C35863              tryptase (EC 3.4.2
    3    1411    93.3    275  2  A35863              tryptase (EC 3.4.2
    4    1405    92.9    275  2  B35863              tryptase (EC 3.4.2
    5    1134    75.0    273  2  A47246              tryptase (EC 3.4.2
    6    1133    74.9    276  2  A38654              mast cell proteina
    7    1131    74.8    275  2  A32410              tryptase (EC 3.4.2
    8  1109.5    73.4    270  2  S56160              mast cell tryptase
    9    1105    73.1    274  2  JC4171              tryptase (EC 3.4.2
   10     999    66.1    237  2  S68702              tryptase (EC 3.4.2
   11     856    56.6    230  2  I48685              mast cell proteina
   12   686.5    45.4    269  2  B32410              mastocytoma protei
   13     516    34.1    343  1  A57014              prostasin (EC 3.4.
   14     468    31.0    625  1  KFHU1               coagulation factor
   15     465    30.8    638  1  KQHUP               plasma kallikrein
   16     461    30.5    638  1  KQMSPL              plasma kallikrein
   17   459.5    30.4    271  1  ELRT2               pancreatic elastas
   18   457.5    30.3    245  1  KYBOB               chymotrypsin (EC 3
   19     454    30.0    455  2  A61545              plasmin (EC 3.4.21
   20     454    30.0    638  1  KQRTPL              plasma kallikrein
   21   451.5    29.9    812  1  PLBO                plasmin (EC 3.4.21
   22     448    29.6    460  2  B61545              plasmin (EC 3.4.21
   23   447.5    29.6    269  2  B26823              pancreatic elastas
   24   440.5    29.1    263  2  A31299              chymotrypsin (EC 3
   25     439    29.0    810  1  PLHU                plasmin (EC 3.4.21
   26   438.5    29.0    263  2  A21195              chymotrypsin (EC 3
   27   438.5    29.0    271  2  A25528              pancreatic elastas
   28     434    28.7    263  1  KYRTB               chymotrypsin (EC 3
   29     434    28.7    810  2  B30848              plasmin (EC 3.4.21
   30   429.5    28.4   4548  1  S00657              apoprotein(a) (EC
   31     429    28.4    266  1  ELPG                pancreatic elastas
   32     429    28.4   1524  2  T30337              polyprotein - Afri
   33   428.5    28.3    367  2  JE0104              testicular serine
   34   426.5    28.2    269  2  C26823              pancreatic elastas
   35   425.5    28.1    812  1  PLMS                plasmin (EC 3.4.21
   36     425    28.1    245  1  KYBOA               chymotrypsin (EC 3
   37     425    28.1    790  1  PLPG                plasmin (EC 3.4.21
   38     424    28.0    246  2  B25528              trypsin (EC 3.4.21
   39   423.5    28.0    436  2  JX0172              acrosin (EC 3.4.21
   40   421.5    27.9    810  2  I46260              plasmin (EC 3.4.21
   41   420.5    27.8    263  2  S47537              chymotrypsin (EC 3
   42   417.5    27.6    270  2  A29934              pancreatic elastas
   43     417    27.6    264  2  I38136              chymotrypsin-like
   44     415    27.4    421  2  S29599              acrosin (EC 3.4.21
   45     414    27.4    786  1  A47547              serine proteinease

                              ALIGNMENTS

RESULT 1
A45754
tryptase (EC 3.4.21.59) alpha precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 08-Sep-1997
C;Accession: A45754; B37193
R;Miller, J.S.; Westin, E.H.; Schwartz, L.B.
J. Clin. Invest. 84, 1188-1195, 1989
A;Title: Cloning and characterization of complementary DNA for human tryptase.
A;Reference number: A45754; MUID:90009311; PMID:2677049
A;Accession: A45754
A;Molecule type: mRNA
A;Residues: 1-274 <MIL>
A;Cross-references: GB:M30038
R;Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A;Title: Cloning and characterization of a second complementary DNA for human tryptase.
A;Reference number: A37193; MUID:90369005; PMID:2203827
A;Accession: B37193
A;Molecule type: mRNA
A;Residues: 1-274 <MI2>
A;Cross-references: GB:M30038
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 2 in having 205-Pro
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-274/Product: tryptase I #status predicted <MAT>
F;31-266/Domain: trypsin homology <TRY>
F;74,120,223/Active site: His, Asp, Ser #status predicted

  Query Match             97.3%;  Score 1470.5;  DB 2;  Length 274;
  Best Local Similarity   97.8%;  Pred. No. 3.5e-123;
  Matches  269;  Conservative    1;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||     ||||||||||||||:||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVN-SGTHLYYQDQLLPVSRIMVHPQFYIIQTGA 119

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 179

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 239

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
```

```
              ||||||||||||||||||||||||||||||||||
Db        240 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 274

RESULT 2
C35863
tryptase (EC 3.4.21.59) III precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-Mar-2003
C;Accession: C35863; E35863; A38893
R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A;Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine pr
A;Reference number: A35863; MUID:90251647; PMID:2187193
A;Accession: C35863
A;Molecule type: mRNA
A;Residues: 9-275 <VAN>
A;Accession: E35863
A;Molecule type: DNA
A;Residues: 1-9 <VA2>
A;Cross-references: GB:M33494; NID:g3927804; PIDN:AAC83172.1; PID:g339977
A;Note: the first nine residues of this sequence are inferred from genomic DNA of trypta
R;Vanderslice, P.
submitted to GenBank, April 1990
A;Reference number: A38893
A;Accession: A38893
A;Molecule type: mRNA
A;Residues: 9-131,'K',132-275 <VA3>
A;Cross-references: GB:M33493; NID:g339984; PIDN:AAA36780.1; PID:g339985
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase I #status predicted <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             94.6%;  Score 1430;  DB 2;  Length 275;
  Best Local Similarity   93.5%;  Pred. No. 1.4e-119;
  Matches  257;  Conservative    8;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||||||||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 3
A35863
tryptase (EC 3.4.21.59) I precursor - human
C;Species: Homo sapiens (man)
C;Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 09-Jul-2004
C;Accession: A35863; D35863; A60939; A39326
R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A;Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine pr
A;Reference number: A35863; MUID:90251647; PMID:2187193
A;Accession: A35863
```

```
A;Molecule type: DNA
A;Residues: 1-275 <VAN>
A;Cross-references: UNIPROT:Q15661; GB:M33494; NID:g3927804; PIDN:AAC83172.1; PID:g33997
A;Accession: D35863
A;Molecule type: mRNA
A;Residues: 1-275 <VA2>
A;Cross-references: GB:M33491
R;Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.
J. Leukoc. Biol. 47, 409-419, 1990
A;Title: Purification of tryptase from a human mast cell line.
A;Reference number: A60939; MUID:90244210; PMID:2110591
A;Accession: A60939
A;Molecule type: protein
A;Residues: 31-38,'P',40-41,'X',43,'T',45-48,'X',50 <BUT>
A;Experimental source: mast cell
A;Note: 44-Gly was also found
R;Cromlish, J.A.; Seidah, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chretien,
J. Biol. Chem. 262, 1363-1373, 1987
A;Title: Human pituitary tryptase: molecular forms, NH-2-terminal sequence, immunocytoche
A;Reference number: A39326; MUID:87109258; PMID:3543004
A;Accession: A39326
A;Molecule type: protein
A;Residues: 31-38 <CRO>
A;Experimental source: pituitary
C;Genetics:
A;Introns: 21/1; 78/2; 177/1; 221/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase I #status experimental <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             93.3%;  Score 1411;  DB 2;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 6.8e-118;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 4
B35863
tryptase (EC 3.4.21.59) II precursor - human
N;Alternate names: tryptase beta
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: B35863; A37193; I59473
R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A;Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine pr
A;Reference number: A35863; MUID:90251647; PMID:2187193
A;Accession: B35863
A;Molecule type: mRNA; DNA
A;Residues: 1-275 <VAN>
```

```
A;Cross-references: UNIPROT:P20231; GB:M33492; NID:g339982; PIDN:AAA36779.1; PID:g339983
A;Note: residues 2-275 are derived from mRNA; residue one was inferred from the genomic
R;Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A;Title: Cloning and characterization of a second complementary DNA for human tryptase.
A;Reference number: A37193; MUID:90369005; PMID:2203827
A;Accession: A37193
A;Molecule type: mRNA
A;Residues: 1-275 <MIL>
A;Cross-references: GB:M37488; NID:g179583; PIDN:AAA51843.1; PID:g179584
R;Blom, T.; Hellman, L.
Scand. J. Immunol. 37, 203-208, 1993
A;Title: Characterization of a tryptase mRNA expressed in the human basophil cell line K
A;Reference number: I59473; MUID:93166209; PMID:8434231
A;Accession: I59473
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-275 <RES>
A;Cross-references: GB:S55551; NID:g265666; PIDN:AAD13876.1; PID:g4261576
A;Experimental source: basophil cell line KU812
C;Genetics:
A;Gene: GDB:TPS1
A;Cross-references: GDB:125890; OMIM:191080
A;Map position: 16pter-16qter
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase I #status predicted <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             92.9%;  Score 1405;  DB 2;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.3e-117;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275

RESULT 5
A47246
tryptase (EC 3.4.21.59) 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A47246
R;McNeil, H.P.; Reynolds, D.S.; Schiller, V.; Ghildyal, N.; Gurley, D.S.; Austen, K.F.;
Proc. Natl. Acad. Sci. U.S.A. 89, 11174-11178, 1992
A;Title: Isolation, characterization, and transcription of the gene encoding mouse mast
A;Reference number: A47246; MUID:93087489; PMID:1454796
A;Accession: A47246
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-273 <MCN>
A;Cross-references: UNIPROT:Q02844; GB:L00653; NID:g200518; PIDN:AAA39992.1; PID:g200519
A;Note: sequence extracted from NCBI backbone (NCBIN:119745, NCBIP:119746)
```

```
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;29-265/Domain: trypsin homology <TRY>

  Query Match             75.0%;  Score 1134;  DB 2;  Length 273;
  Best Local Similarity   74.4%;  Pred. No. 2.8e-93;
  Matches  203;  Conservative   23;  Mismatches   45;  Indels    2;  Gaps    1;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              || |||| ||:|:|  :||| |  |: : ||||||||  :|||||||||  | |||||||
Db          1 MLKLLLLTLPLLSSLVHAAPGP--AMTREGIVGGQEAHGNKWPWQVSLRANDTYWMHFCG 58

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||| |   :|||||:|:||| | |: ||:|| || |||:| ||
Db         59 GSLIHPQWVLTAAHCVGPDVADPNKVRVQLRKQYLYYHDHLMTVSQIITHPDFYIVQDGA 118

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              |||||:|  |||||  || | ||||||||| |  |||||||::||   ||||||||:|:|
Db        119 DIALLKLTNPVNISDYVHPVPLPPASETFPSGTLCWVTGWGNIDNGVNLPPPFPLKEVQV 178

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||:|||:|| ||| |  |||:| |:|||||||||   |||:|||||||||||  ||||||
Db        179 PIIENHLCDLKYHKGLITGDNVHIVRDDMLCAGNEGHDSCQGDSGGPLVCKVEDTWLQAG 238

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              |||| ||||||||||||||||||||||||||||
Db        239 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPK 271

RESULT 6
A38654
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Feb-1992 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: A38654; B38654; D35646; I59478
R;Reynolds, D.S.; Gurley, D.S.; Austen, K.F.; Serafin, W.E.
J. Biol. Chem. 266, 3847-3853, 1991
A;Title: Cloning of the cDNA and gene of mouse mast cell protease-6. Transcription by prc
A;Reference number: A38654; MUID:91139682; PMID:1995638
A;Accession: A38654
A;Molecule type: DNA
A;Residues: 1-276 <REY>
A;Cross-references: UNIPROT:P21845; GB:M57625; NID:g200506; PIDN:AAA39987.1; PID:g200507
A;Note: the authors translated the codon CGC for residue 24 as Ala, GAG for residue 37 as
s Gly, GAG for residue 148 as Gly, GAG for residue 168 as Gly, and GAA for 185 as Gly
A;Accession: B38654
A;Molecule type: mRNA
A;Residues: 1-276 <RE2>
A;Cross-references: GB:M57626; NID:g200508; PIDN:AAA39988.1; PID:g200509
R;Reynolds, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.E.
Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990
A;Title: Different mouse mast cell populations express various combinations of at least s
A;Reference number: A35646; MUID:90222202; PMID:2326280
A;Accession: D35646
A;Molecule type: protein
A;Residues: 32-54 <RE3>
R;Huang, R.; Abrink, M.; Gobl, A.E.; Nilsson, G.; Aveskogh, M.; Larsson, L.G.; Nilsson, I
Scand. J. Immunol. 38, 359-367, 1993
A;Title: Expression of a mast cell tryptase in the human monocytic cell lines U-937 and m
A;Reference number: I59478; MUID:94023807; PMID:8210998
A;Accession: I59478
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-276 <RES>
A;Cross-references: GB:L31853; NID:g473480; PIDN:AAA39725.1; PID:g473481
C;Genetics:
A;Gene: MMCP-6
A;Introns: 24/1; 79/2; 168/1; 222/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-31/Domain: activation peptide #status predicted <ACT>
F;32-276/Product: mast cell proteinase 6 #status experimental <MAT>
```

```
F;32-268/Domain: trypsin homology <TRY>
F;75,122,225/Active site: His, Asp, Ser #status predicted

  Query Match             74.9%;  Score 1133;  DB 2;  Length 276;
  Best Local Similarity   73.3%;  Pred. No. 3.5e-93;
  Matches  198;  Conservative   28;  Mismatches   42;  Indels    2;  Gaps    1;

Qy          4 LLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSL 63
              ||| || :|||  |:|| |  | |: ||||| ||  |||||||||| :  ||:|||||||
Db          7 LLLWALSLLASLVYSAPRP--ANQRVGIVGGHEASESKWPWQVSLRFKLNYWIHFCGGSL 64

Qy         64 IHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIA 123
              ||||||||||||:|| :|     |||||||:||| |||| ::||:||| :|  : |||:|
Db         65 IHPQWVLTAAHCVGPHIKSPQLFRVQLREQYLYYGDQLLSLNRIVVHPHYYTAEGGADVA 124

Qy        124 LLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIM 183
              ||||| |||:|: :| : |||||||||||  |||||||||:||||||||||:|||||||||:
Db        125 LLELEVPVNVSTHIHPISLPPASETFPPGTSCWVTGWGDIDNDEPLPPPYPLKQVKVPIV 184

Qy        184 ENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVS 243
              || :|| ||| | |||||  |: | ||||||::||||:|||||||||||| ||||||||||
Db        185 ENSLCDRKYHTGLYTGDDFPIVHDGMLCAGNTRRDSCQGDSGGPLVCKVKGTWLQAGVVS 244

Qy        244 WDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              | |||||||:||||||||||||||| |||:
Db        245 WGEGCAQPNKPGIYTRVTYYLDWIHRYVPE 274

RESULT 7
A32410
tryptase (EC 3.4.21.59) precursor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 09-Jul-2004
C;Accession: A32410
R;Vanderslice, P.; Craik, C.S.; Nadel, J.A.; Caughey, G.H.
Biochemistry 28, 4148-4155, 1989
A;Title: Molecular cloning of dog mast cell tryptase and a related protease: structural
A;Reference number: A32410; MUID:89352460; PMID:2504277
A;Accession: A32410
A;Molecule type: mRNA
A;Residues: 1-275 <VAN>
A;Cross-references: UNIPROT:P15944; GB:M24664; NID:g163982; PIDN:AAA30854.1; PID:g163983
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase #status predicted <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             74.8%;  Score 1131;  DB 2;  Length 275;
  Best Local Similarity   72.7%;  Pred. No. 5.3e-93;
  Matches  200;  Conservative   30;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              | | |:||| :| |    :||| ||||: |||||:||| |||||||||||::  :|| | ||
Db          1 MPSPLVLALALLGSLVPVSPAPGQALQRVGIVGGREAPGSKWPWQVSLRLKGQYWRHICG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:||:|     :||||||||||||| ||||:||::|| :|  : ||
Db         61 GSLIHPQWVLTAAHCVGPNVVCPEEIRVQLREQHLYYQDHLLPVNRIVMHPNYYTPENGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||:|||:|: |  | |||| :||| | |||||||||| :  |||||||||||||
Db        121 DIALLELEDPVNVSAHVQPVTLPPALQTFPTGTPCWVTGWGDVHSGTPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||:|| :|| :||||  ||| |||:|:||||||||: |||:|||||||||:| | |||||
Db        181 PIVENSMCDVQYHLGLSTGDGVRIVREDMLCAGNSKSDSCQGDSGGPLVCRVRGVWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||| ||||||| ||||:|
Db        241 VVSWGEGCAQPNRPGIYTRVAYYLDWIHQYVPKEP 275
```

```
RESULT 8
S56160
mast cell tryptase precursor - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S56160
R;Murakumo, Y.; Ide, H.; Itoh, H.; Tomita, M.; Kobayashi, T.; Maruyama, H.; Horii, Y.; Na
Biochem. J. 309, 921-926, 1995
A;Title: Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil, Meriones ur
A;Reference number: S56160; MUID:95366971; PMID:7639711
A;Accession: S56160
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-270 <MUR>
A;Cross-references: UNIPROT:P50342; EMBL:D31789; NID:g517122; PIDN:BAA06598.1; PID:g51712
C;Superfamily: trypsin; trypsin homology
F;26-262/Domain: trypsin homology <TRY>

  Query Match             73.4%;  Score 1109.5;  DB 2;  Length 270;
  Best Local Similarity   73.3%;  Pred. No. 4.2e-91;
  Matches  200;  Conservative   22;  Mismatches   46;  Indels    5;  Gaps    2;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              || ||||||| | |  : :|      |: ||||||||| :|||||||||   : || ||||
Db          1 MLKLLLLALP-LFSLMHRSP----LCQEWGIVGGQEAPGNKWPWQVSLRANETYWRHFCG 55

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|| : |   :|||||:|:||| | || ||||| || ||  | ||
Db         56 GSLIHPQWVLTAAHCVGPTIADPNKVRVQLRKQYLYYHDHLLAVSRIITHPTFYATQNGA 115

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              |||||||: ||||||| || | ||||||||| |  |||||||::|||  ||||||||:|:|
Db        116 DIALLELKNPVNISSHVHPVSLPPASETFPSGTLCWVTGWGNIDNDVSLPPPFPLKEVQV 175

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              |::|| :|| ||| | ||||:: |:|||||||||   |||:|||||||||||||||||||
Db        176 PVVENQLCDLKYHKGVYTGDNIHIVRDDMLCAGNEGHDSCQGDSGGPLVCKVNGTWLQAG 235

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              |||| |||| |||||||||||||||||| ||||
Db        236 VVSWGEGCALPNRPGIYTRVTYYLDWIHRYVPK 268

RESULT 9
JC4171
tryptase (EC 3.4.21.59) precursor - rat
N;Alternate names: mast cell tryptase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Aug-1995 #sequence_revision 27-Oct-1995 #text_change 09-Jul-2004
C;Accession: JC4171
R;Ide, H.; Itoh, H.; Tomita, M.; Murakumo, Y.; Kobayashi, T.; Maruyama, H.; Osada, Y.; Na
J. Biochem. 118, 210-215, 1995
A;Title: cDNA sequencing and expression of rat mast cell tryptase.
A;Reference number: JC4171; MUID:96015171; PMID:8537314
A;Accession: JC4171
A;Molecule type: mRNA
A;Residues: 1-274 <IDE>
A;Cross-references: UNIPROT:P50343; DDBJ:D38455; NID:g556555; PIDN:BAA07486.1; PID:g55655
C;Comment: This enzyme is basically specific for a connective tissue mast cell, it is up-
einase inhibitors.
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; mast cell; serine proteinase; zymogen
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-29/Domain: activation peptide #status predicted <ACT>
F;30-274/Product: mast cell tryptase #status predicted <MAT>
F;30-266/Domain: trypsin homology <TRY>
F;73,120,223/Active site: His, Asp, Ser #status predicted
F;131/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             73.1%;  Score 1105;  DB 2;  Length 274;
  Best Local Similarity   72.0%;  Pred. No. 1.1e-90;
```

```
  Matches  195;  Conservative   31;  Mismatches   43;  Indels    2;  Gaps    1;

Qy          4 LLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSL 63
              ||||||  |||  :||| ||:  |: |||||:||  ||||||||| :  :||||||||||
Db          5 LLLLALSPLASLVHAAPCPVK--QRVGIVGGREASESKWPWQVSLRFKFSFWMHFCGGSL 62

Qy         64 IHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIA 123
              ||||||||||||:|  :|     |||||||:||| |||| |:| :||| :| :: |||||
Db         63 IHPQWVLTAAHCVGLHIKSPELFRVQLREQYLYYADQLLTVNRTVVHPHYYTVEDGADIA 122

Qy        124 LLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIM 183
              ||||| |||:|: :| : ||||||||| |  ||||||||:|:|||| ||:|||||||||:
Db        123 LLELEIPVNVSTHIHPISLPPASETFPSGTSCWVTGWGDIDSDEPLLPPYPLKQVKVPIV 182

Qy        184 ENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVS 243
              || :|| ||| | |||||| |::| ||||||:: |||:||||||||||| ||||||||||
Db        183 ENSLCDRKYHTGLYTGDDVPIVQDGMLCAGNTRSDSCQGDSGGPLVCKVKGTWLQAGVVS 242

Qy        244 WDEGCAQPNRPGIYTRVTYYLDWIHHYVPKK 274
              | ||||: ||||||||||||||||| |||::
Db        243 WGEGCAEANRPGIYTRVTYYLDWIHRYVPQR 273

RESULT 10
S68702
tryptase (EC 3.4.21.59) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: S68702; S68676
R;Pallaoro, M.; Gambacurta, A.; Fiorucci, L.; Mignogna, G.; Barra, D.; Ascoli, F.
Eur. J. Biochem. 237, 100-105, 1996
A;Title: cDNA cloning and primary structure of tryptase from bovine mast cells, and evid
A;Reference number: S68676; MUID:96203914; PMID:8620861
A;Accession: S68702
A;Molecule type: mRNA
A;Residues: 1-237 <PAL>
A;Cross-references: UNIPROT:Q29464; EMBL:X94982; NID:g1332446; PIDN:CAA64438.1; PID:g133
A;Accession: S68676
A;Molecule type: protein
A;Residues: 'IVGGQEAP',1-8;61-74;90-97;126-148;162-190;208-222;227-235 <PAW>
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; serine proteinase; zymogen
F;1-229/Domain: trypsin homology (fragment) <TRY>
F;36,83,177/Active site: His, Asp, Ser #status predicted
F;94,195/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             66.1%;  Score 999;  DB 2;  Length 237;
  Best Local Similarity   74.6%;  Pred. No. 2.4e-81;
  Matches  176;  Conservative   22;  Mismatches   38;  Indels    0;  Gaps    0;

Qy         40 SKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQD 99
              |:|||||||||  ||| | |||||||||||||||||:||:|   :  |||||||||||||
Db          2 SQWPWQVSLRVSRRYWRHHCGGSLIHPQWVLTAAHCVGPEVHGPSYFRVQLREQHLYYQD 61

Qy        100 QLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTG 159
              ||||:|||| ||  | :: |||||||||:: ||||  |  | ||| |||||||  |||||
Db         62 QLLPISRIIPHPNCYSVKNGADIALLELDKLVNISWHVQPVTLPPESETFPPGTQCWVTG 121

Qy        160 WGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDS 219
              ||:|||   |||||||||||||::|| :|| ||| |  |||:| |:|:||||||:| |:
Db        122 WGNVDNGRRLPPPFPLKQVKVPVVENSVCDRKYHSGLSTGDNVPIVREDMLCAGDSGRNF 181

Qy        220 CKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |:||||||||||||||||||||||| :|||:||||||||||| |||||| |||: |
Db        182 CQGDSGGPLVCKVNGTWLQAGVVSWGDGCAKPNRPGIYTRVTSYLDWIHQYVPQGP 237

RESULT 11
I48685
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48685; S43172
R;Huang, R.; Hellman, L.
Immunogenetics 40, 397-414, 1994
A;Title: Genes for mast-cell serine protease and their molecular evolution.
A;Reference number: I48684; MUID:95048582; PMID:7959952
A;Accession: I48685
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-230 <RES>
A;Cross-references: UNIPROT:P21845; EMBL:X78542; NID:g468809; PIDN:CAA55288.1; PID:g4688
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;32-230/Domain: trypsin homology #status atypical <TRY>

  Query Match             56.6%;  Score 856;  DB 2;  Length 230;
  Best Local Similarity   69.3%;  Pred. No. 1.2e-68;
  Matches  151;  Conservative   26;  Mismatches   39;  Indels    2;  Gaps    1;

Qy          4 LLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSL 63
              ||| || :|||  |:|| |  | |: ||||| ||  ||||||||| :  ||:||||||||
Db          7 LLLWALSLLASLVYSAPRP--ANQRVGIVGGHEASESKWPWQVSLRFKLNYWIHFCGGSL 64

Qy         64 IHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIA 123
              ||||||||||||:|| :|     |||||||:||| |||| ::||:||| :|  : |||:|
Db         65 IHPQWVLTAAHCVGPHIKSPQLFRVQLREQYLYYGDQLLSLNRIVVHPHYYTAEGGADVA 124

Qy        124 LLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIM 183
              ||||| |||:|: :| : |||||||||||  ||||||||:|||||||||:|||||||||:
Db        125 LLELEVPVNVSTHIHPISLPPASETFPPGTSCWVTGWGDIDNDEPLPPPYPLKQVKVPIV 184

Qy        184 ENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCK 221
              || :|| ||| | |||||  |: | ||||||::||||:
Db        185 ENSLCDRKYHTGLYTGDDFPIVHDGMLCAGNTRRDSCQ 222

RESULT 12
B32410
mastocytoma proteinase (EC 3.4.21.-) precursor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 12-Oct-1989 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: B32410
R;Vanderslice, P.; Craik, C.S.; Nadel, J.A.; Caughey, G.H.
Biochemistry 28, 4148-4155, 1989
A;Title: Molecular cloning of dog mast cell tryptase and a related protease: structural e
A;Reference number: A32410; MUID:89352460; PMID:2504277
A;Accession: B32410
A;Molecule type: mRNA
A;Residues: 1-269 <VAN>
A;Cross-references: UNIPROT:P19236; GB:M24665; NID:g163984; PID:g163985; GB:J02862
A;Note: the authors translated the codon ACG for residue 114 as Ser
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-269/Product: mastocytoma proteinase #status predicted <MAT>
F;20-259/Domain: trypsin homology <TRY>
F;66,116,217/Active site: His, Asp, Ser #status predicted

  Query Match             45.4%;  Score 686.5;  DB 2;  Length 269;
  Best Local Similarity   52.0%;  Pred. No. 1.7e-53;
  Matches  132;  Conservative   34;  Mismatches   79;  Indels    9;  Gaps    4;

Qy         28 QAGIVGGQEAPRSKWPWQVSLR---VRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLA 84
              : ||||| : |  ::|||||||   :    | | |||||||||||||||||:  :  : |
Db         17 KVGIVGGCKVPARRYPWQVSLRFHGMGSGQWQHICGGSLIHPQWVLTAAHCVELEGLEAA 76

Qy         85 TLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTG---ADIALLELEEPVNISSRVHTVM 141
              |||||: :  ||  |||  |: || || | :   |   ||||||:|| |: :|  |: |
Db         77 TLRVQVGQLRLYDHDQLCNVTEIIRHPNFNMSWYGWDTADIALLKLEAPLTLSEDVNLVS 136

Qy        142 LPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDD 201
              ||  |   |||| ||||||||: :  |||||: |::|:|||:  |  |:  |        |
Db        137 LPSPSLIVPPGMLCWVTGWGDIADHTPLPPPYHLQEVEVPIVGNRECNCHYQ--TILEQD 194

Qy        202 VRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVT 261
```

```
                  :|: ||||||:   |||: ||||||||:   ||:| |||||  ||   | ||:| |||
Db          195 DEVIKQDMLCAGSEGHDSCQMDSGGPLVCRWKCTWIQVGVVSWGYGCGY-NLPGVYARVT 253

Qy          262 YYLDWIHHYVPKKP 275
                 |: ||| ::|  |
Db          254 SYVSWIHQHIPLSP 267


RESULT 13
A57014
prostasin (EC 3.4.21.-) precursor - human
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C;Accession: A57014; A54866
R;Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 270, 13483-13489, 1995
A;Title: Molecular cloning, tissue-specific expression, and cellular localization of hum
A;Reference number: A57014; MUID:95286644; PMID:7768952
A;Accession: A57014
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-343 <RES>
A;Cross-references: UNIPROT:Q16651; GB:L41351; NID:g862304; PIDN:AAC41759.1; PID:g862305
A;Experimental source: prostate
A;Note: parts of this sequence were determined by protein sequencing
R;Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 269, 18843-18848, 1994
A;Title: Prostasin is a novel human serine proteinase from seminal fluid. Purification,
A;Reference number: A54866; MUID:94308140; PMID:8034638
A;Accession: A54866
A;Molecule type: protein
A;Residues: 45-64 <YUA>
C;Genetics:
A;Gene: GDB:PRSS8
A;Cross-references: GDB:676446; OMIM:600823
A;Map position: 16p11.2-16p11.2
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein
F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-44,45-343/Product: prostasin #status predicted <MAT>
F;33-44/Domain: prostasin light chain #status predicted <CHL>
F;45-343/Domain: prostasin heavy chain #status predicted <CHH>
F;45-281/Domain: trypsin homology <TRY>
F;323-341/Domain: transmembrane #status predicted <TMM1>
F;37-154,70-86,168-244,201-223,234-262/Disulfide bonds: #status predicted
F;85,134,238/Active site: His, Asp, Ser #status predicted
F;159/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             34.1%;  Score 516;  DB 1;  Length 343;
  Best Local Similarity   40.0%;  Pred. No. 3.3e-38;
  Matches  110;  Conservative   42;  Mismatches  109;  Indels   14;  Gaps    8;

Qy            4 LLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSL 63
                || | |    : |  | ||     || | ||  |   :||||||:       :| |||||
Db           18 LLYLGLLRSGTGAEGAEAPCGVAPQARITGGSSAVAGQWPWQVSITYEG---VHVCGGSL 74

Qy           64 IHPQWVLTAAHCLGPDVKDLATLRVQLREQHL--YYQD-QLLPVSRIIVHPQFYIIQTGA 120
                :  ||||:||||  |         |:|    |  | :| ::  :  || || :  :| |:
Db           75 VSEQWVLSAAHCF-PSEHHKEAYEVKLGAHQLDSYSEDAKVSTLKDIIPHPSY--LQEGS 131

Qy          121 --DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQV 178
                  |||||:|  |:  |  :  : || |: :|| |: | ||||| |     |  | ||:|:
Db          132 QGDIALLQLSRPITFSRYIRPICLPAANASFPNGLHCTVTGWGHVAPSVSLLTPKPLQQL 191

Qy          179 KVPIMENHICDAKYHLGAYTGDDVRIIRDDMLCAG--NSQRDSCKGDSGGPLVCKVNGTW 236
                :||::    |:  |:: |   ::   :::||:|||     :|:|:||||||| | | | |
Db          192 EVPLISRETCNCLYNIDA-KPEEPHFVQEDMVCAGYVEGGKDACQGDSGGPLSCPVEGLW 250

Qy          237 LQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHYV 271
                   |:||| : |   ||||:||  : |  ||   |
Db          251 YLTGIVSWGDACGARNRPGVYTLASSYASWIQSKV 285
```

```
RESULT 14
KFHU1
coagulation factor XIa (EC 3.4.21.27) precursor [validated] - human
N;Alternate names: antihemophilic factor C; plasma thromboplastin antecedent
C;Species: Homo sapiens (man)
C;Date: 13-Aug-1986 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C;Accession: A27431; A00920; A37940
R;Asakai, R.; Davie, E.W.; Chung, D.W.
Biochemistry 26, 7221-7228, 1987
A;Title: Organization of the gene for human factor XI.
A;Reference number: A27431; MUID:88107663; PMID:2827746
A;Accession: A27431
A;Molecule type: DNA
A;Residues: 1-625 <ASA>
A;Cross-references: UNIPROT:P03951; GB:M18295
A;Note: the sequence shown follows the authors' translation
R;Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
Biochemistry 25, 2417-2424, 1986
A;Title: Amino acid sequence of human factor XI, a blood coagulation factor with four tan
A;Reference number: A00920; MUID:86243360; PMID:3636155
A;Accession: A00920
A;Molecule type: mRNA
A;Residues: 1-625 <FUJ>
A;Cross-references: GB:M13142; NID:g182832; PIDN:AAA52487.1; PID:g182833
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2056-2060, 1991
A;Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
A;Reference number: A37940; MUID:91152017; PMID:1998667
A;Accession: A37940
A;Molecule type: protein
A;Residues: 28-33;35-49,'X',51-55,'X',57-63;70-75,'X',77-79;107-109,'X',111-112;132-139;'
;280-282,'X',284;285-297;313-316,'X',318-319;320-326;'X',328-330;'X',347-349;373,'X',375;
C;Comment: The proenzyme consists of two identical chains linked by one or more disulfide
he active site, and a heavy chain, which associates with high molecular weight (HMW) kini
C;Genetics:
A;Gene: GDB:F11
A;Cross-references: GDB:119891; OMIM:264900
A;Map position: 4q35-4q35
A;Introns: 19/1; 73/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 526
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor IX
A;Pathway: blood coagulation intrinsic pathway
C;Superfamily: coagulation factor XI; trypsin homology
C;Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydrol
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-387/Product: coagulation factor XIa heavy chain #status experimental <HCH>
F;19-108/Domain: apple repeat <AP1>
F;109-198/Domain: apple repeat <AP2>
F;199-288/Domain: apple repeat <AP3>
F;290-379/Domain: apple repeat <AP4>
F;388-625/Product: coagulation factor XIa light chain #status experimental <LCH>
F;388-618/Domain: trypsin homology <TRY>
F;20-103,514-581,571-599/Disulfide bonds: #status predicted
F;29/Disulfide bonds: interchain #status experimental
F;46-76,50-56,110-193,136-165,140-146,200-283,226-255,230-236,291-374,317-346,321-327,380
F;90,126,353,450/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;339/Disulfide bonds: interchain #status predicted
F;387-388/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;431,480,575/Active site: His, Asp, Ser #status predicted
F;491/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             31.0%;  Score 468;  DB 1;  Length 625;
  Best Local Similarity   39.5%;  Pred. No. 1.2e-33;
  Matches   98;  Conservative   31;  Mismatches   91;  Indels   28;  Gaps    6;

Qy           31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQ- 89
                ||||  : | :|||||:|        | ||||:|  ||:||||||    |:    |||
Db          388 IVGGTASVRGEWPWQVTLHTTSPTQRHLCGGSIIGNQWILTAAHCF-YGVESPKILRVYS 446

Qy           90 --LREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASE 147
                  | :  :        |  ||:| |: : ::| |||||:||  || :     : ||   :
Db          447 GILNQSEIKEDTSFFGVQEIIIHDQYKMAESGYDIALLKLETTVNYTDSQRPICLPSKGD 506
```

```
Qy        148 TFPPGMPCWVTGWG------DVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDD 201
                     |||||||       : |         |:: |:|:: |  |  :|
Db        507 RNVIYTDCWVTGWGYRKLRDKIQN--------TLQKAKIPLVTNEECQKRYR-------- 550

Qy        202 VRIIRDDMLCAGNSQ--RDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTR 259
                 |   |:|||  :  :|:||||||||| || |  |   |: || |||||  |||:||
Db        551 GHKITHKMICAGYREGGKDACKGDSGGPLSCKHNEVWHLVGITSWGEGCAQRERPGVYTN 610

Qy        260 VTYYLDWI 267
              |  |:|||
Db        611 VVEYVDWI 618

RESULT 15
KQHUP
plasma kallikrein (EC 3.4.21.34) precursor - human
N;Alternate names: kininogenin; plasma prekallikrein
C;Species: Homo sapiens (man)
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 09-Jul-2004
C;Accession: A00921; A37939
R;Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
Biochemistry 25, 2410-2417, 1986
A;Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four t
A;Reference number: A00921; MUID:86243359; PMID:3521732
A;Accession: A00921
A;Molecule type: mRNA
A;Residues: 1-638 <CHU>
A;Cross-references: UNIPROT:P03952; GB:M13143; NID:g190262; PIDN:AAA60153.1; PID:g190263
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2050-2056, 1991
A;Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of 
A;Reference number: A37939; MUID:91152016; PMID:1998666
A;Accession: A37939
A;Molecule type: protein
A;Residues: 20-27;40-46,'X',48,'H';50,'X',52-70,'H';75-76,'X',78-80;103-113;131-140;141-
;260-283,'X',285;287-291,'X',293-295;314-317,'X',319-320;321-324;'X',329-333;334-339,'X'
525;538-551;562,'X',564-567;573,'X',575-576;578-583,'X',585;592-604 <MCM>
C;Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
C;Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li
 are linked by one or more disulfide bonds.
C;Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal r
inogen and may also play a role in the renin-angiotensin system by converting prorenin i
C;Genetics:
A;Gene: GDB:KLK3
A;Cross-references: GDB:127575; OMIM:229000
A;Map position: 4q35-4q35
C;Superfamily: coagulation factor XI; trypsin homology
C;Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-638/Product: plasma kallikrein #status predicted <MAT>
F;20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>
F;20-109/Domain: apple repeat <AP1>
F;110-199/Domain: apple repeat <AP2>
F;200-289/Domain: apple repeat <AP3>
F;291-380/Domain: apple repeat <AP4>
F;391-638/Domain: plasma kallikrein light chain #status predicted <LCH>
F;391-621/Domain: trypsin homology <TRY>
F;21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383
F;127,308,396,453,494/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;318-347,340-345/Disulfide bonds: #status predicted
F;390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted
F;434,483,578/Active site: His, Asp, Ser #status predicted

  Query Match             30.8%;  Score 465;  DB 1;  Length 638;
  Best Local Similarity   39.0%;  Pred. No. 2.3e-33;
  Matches   97;  Conservative   41;  Mismatches   81;  Indels   30;  Gaps    8;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCL-GPDVKDLATLRVQ 89
              ||||  :   :|||||||:|:     | ||||||  |||||||||  |  ::|:   |:
Db        391 IVGGTNSSWGEWPWQVSLQVKLTAQRHLCGGSLIGHQWVLTAAHCFDGLPLQDV--WRIY 448

Qy         90 ---LREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPAS 146
                 |    :        :  ||:|  : : :   ||||::|: |:| :     : ||

Db        449 SGILNLSDITKDTPFSQIKEIIIHQNYKVSEGNHDIALIKLQAPLNYTEFQKPICLPSKG 508

Qy        147 ETFPPGMPCWVTGW------GDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGD 200
              :|      ||||||      |:: |         |::| :|:: |  |  :|
Db        509 DTSTIYTNCWVTGWGFSKEKGEIQN--------ILQKVNIPLVTNEECQKRY-------Q 553

Qy        201 DVRIIRDDMLCAGNSQ--RDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYT 258
              | :| :  |:|||  :  :|:|||||||||||| || |   |: || ||||:  :||:||
Db        554 DYKITQ-RMVCAGYKEGGKDACKGDSGGPLVCKHNGMWRLVGITSWGEGCARREQPGVYT 612

Qy        259 RVTYYLDWI 267
              :|  |:|||
Db        613 KVAEYMDWI 621

Search completed: August 27, 2005, 23:30:12
Job time : 58 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         August 27, 2005, 23:13:05 ; Search time 128 Seconds
                                            (without alignments)
                                            1100.170 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:      1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       UniProt_03:*
                1:  uniprot_sprot:*
                2:  uniprot_trembl:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                          Description
-------------------------------------------------------------------------------
    1     1505    99.5    275  1  TRYA_HUMAN                  P15157 homo sapien
    2     1494    98.8    275  2  Q86TM8                      Q86tm8 homo sapien
    3     1494    98.8    276  2  Q86UA5                      Q86ua5 homo sapien
    4     1414    93.5    275  2  Q6B051                      Q6b051 homo sapien
    5     1411    93.3    275  1  TRB1_HUMAN                  Q15661 homo sapien
    6     1405    92.9    275  1  TRB2_HUMAN                  P20231 homo sapien
    7     1405    92.9    282  2  Q6NZY1                      Q6nzy1 homo sapien
    8     1403    92.8    275  2  Q6FHB8                      Q6fhb8 homo sapien
    9     1401    92.7    275  2  Q96RZ6                      Q96rz6 homo sapien
   10     1156    76.5    275  2  Q7YS62                      Q7ys62 equus cabal
   11     1134    75.0    273  1  MCT7_MOUSE                  Q02844 mus musculu
   12     1133    74.9    276  1  MCT6_MOUSE                  P21845 mus musculu
   13     1131    74.8    275  1  TRYT_CANFA                  P15944 canis famil
   14     1123    74.3    273  2  Q921N4                      Q921n4 mus musculu
   15     1120    74.1    233  2  Q96RZ7                      Q96rz7 homo sapien
   16     1118    73.9    273  1  TRYT_SHEEP                  Q9xsm2 ovis aries
   17     1110    73.4    242  2  Q6NTB8                      Q6ntb8 homo sapien
   18   1109.5    73.4    270  1  TRYT_MERUN                  P50342 meriones un
   19     1109    73.3    275  1  TRYT_PIG                    Q9n2d1 sus scrofa
   20     1105    73.1    273  1  MCT7_RAT                    P27435 rattus norv
   21     1105    73.1    273  2  Q6P6W8                      Q6p6w8 rattus norv
   22     1105    73.1    274  1  MCT6_RAT                    P50343 rattus norv
   23     1103    72.9    273  2  Q9XSM1                      Q9xsm1 ovis aries
   24     1100    72.8    235  1  TRYD_HUMAN                  Q9bzj3 homo sapien
   25     1098    72.6    271  2  Q8HYJ2                      Q8hyj2 bos taurus
   26      999    66.1    237  2  Q29464                      Q29464 bos taurus
   27    717.5    47.5    280  2  Q66NX6                      Q66nx6 canis famil
   28    716.5    47.4    280  2  Q66NX5                      Q66nx5 canis famil
   29      697    46.1    318  2  Q80UR4                      Q80ur4 mus musculu
   30      692    45.8    314  2  Q6RUT2                      Q6rut2 mus musculu
   31      689    45.6    277  2  Q8SQ44                      Q8sq44 sus scrofa
```

```
   32    686.5    45.4    269  1  TRYM_CANFA                  P19236 canis famil
   33      649    42.9    279  2  Q99MS4                      Q99ms4 mus musculu
   34    632.5    41.8    321  1  TRYG_HUMAN                  Q9nrr2 homo sapien
   35    628.5    41.6    321  2  Q96RZ8                      Q96rz8 homo sapien
   36      621    41.1    274  2  Q924N9                      Q924n9 mus musculu
   37    603.5    39.9    311  1  TRYG_MOUSE                  Q9qul7 mus musculu
   38    600.5    39.7    277  2  Q80WM7                      Q80wm7 mus musculu
   39    599.5    39.6    328  2  Q6BEA2                      Q6bea2 rattus norv
   40      597    39.5    321  2  Q6IE60                      Q6ie60 rattus norv
   41    592.5    39.2    328  2  Q8BJR6                      Q8bjr6 mus musculu
   42      578    38.2    311  2  Q80XZ3                      Q80xz3 rattus norv
   43      576    38.1    330  2  Q6NVR7                      Q6nvr7 xenopus tro
   44    572.5    37.9    290  1  PR27_HUMAN                  Q9bqr3 homo sapien
   45    561.5    37.1    719  2  Q6DJ90                      Q6dj90 xenopus tro


                                ALIGNMENTS


RESULT 1
TRYA_HUMAN
ID   TRYA_HUMAN     STANDARD;      PRT;   275 AA.
AC   P15157; Q9H2Y5; Q9UQI1;
DT   01-APR-1990 (Rel. 14, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-JAN-2005 (Rel. 46, Last annotation update)
DE   Alpha-tryptase precursor (EC 3.4.21.59) (Tryptase 1).
GN   Name=TPS1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Lung;
RX   MEDLINE=90009311; PubMed=2677049;
RA   Miller J.S., Westin E.H., Schwartz L.B.;
RT   "Cloning and characterization of complementary DNA for human
RT   tryptase.";
RL   J. Clin. Invest. 84:1188-1195(1989).
RN   [2]
RP   REVISIONS TO 89-93 AND 108.
RA   Schwartz L.B.;
RL   Submitted (MAR-1990) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=99121069; PubMed=9920877; DOI=10.1074/jbc.274.6.3355;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Lung;
RA   Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA   Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA   Hunt J.E.;
RT   "Molecular cloning and characterization of novel human tryptase cDNAs
RT   and splicing variants.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 31-50, AND PITUITARY, SEQUENCE OF 31-38.
RC   TISSUE=Lung;
RX   MEDLINE=87109258; PubMed=3543004;
RA   Cromlish J.A., Seidah N.G., Marcinkiewcz M., Hamelin J., Johnson D.A.,
RA   Chretein M.;
RT   "Human pituitary tryptase: molecular forms, NH2-terminal sequence,
RT   immunocytochemical localization, and specificity with prohormone and
RT   fluorogenic substrates.";
RL   J. Biol. Chem. 262:1363-1373(1987).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
```

```
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P15157-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P15157-2; Sequence=VSP_005374;
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Tryptase
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M30038; AAA86934.1; -.
DR   EMBL; AF098328; AAD17846.1; -.
DR   EMBL; AF206665; AAG35695.1; -.
DR   EMBL; AF206666; AAG35696.1; -.
DR   PDB; 1LTO; X-ray; A/B/C/D=31-275.
DR   MEROPS; S01.143; -.
DR   Genew; HGNC:12018; TPS1.
DR   MIM; 191080; -.
DR   GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR   GO; GO:0006952; P:defense response; TAS.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   3D-structure; Alternative splicing; Direct protein sequencing;
KW   Glycoprotein; Hydrolase; Polymorphism; Serine protease; Signal;
KW   Zymogen.
FT   SIGNAL        1     18       Potential.
FT   PROPEP       19     30       Activation peptide.
FT   CHAIN        31    275       Alpha-tryptase.
FT   ACT_SITE     74     74       Charge relay system.
FT   ACT_SITE    121    121       Charge relay system.
FT   ACT_SITE    224    224       Charge relay system.
FT   DISULFID     59     75
FT   DISULFID    155    230
FT   DISULFID    188    211
FT   DISULFID    220    248
FT   CARBOHYD    132    132       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC     79     87       Missing (in isoform 2).
FT                                /FTId=VSP_005374.
FT   VARIANT      15     15       R -> P (in alpha-II).
FT                                /FTId=VAR_012102.
FT   VARIANT     221    221       K -> Q (in alpha-II; dbSNP:1137382).
FT                                /FTId=VAR_012103.
FT   CONFLICT    215    216       TR -> SQ (in Ref. 1).
SQ   SEQUENCE   275 AA;  30772 MW;  B9BAC4BBCB91CE75 CRC64;

  Query Match             99.5%;  Score 1505;  DB 1;  Length 275;
  Best Local Similarity   99.3%;  Pred. No. 1.6e-122;
  Matches  273;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              |||||||||||||||||||||||||||||||||||::|||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 2
Q86TM8
ID   Q86TM8      PRELIMINARY;      PRT;   275 AA.
AC   Q86TM8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   TPS1 protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC028059; AAH28059.1; -.
DR   HSSP; P15157; 1LTO.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
```

```
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   275 AA;  30713 MW;  189C3EBBCB473FA3 CRC64;

  Query Match             98.8%;  Score 1494;  DB 2;  Length 275;
  Best Local Similarity   98.5%;  Pred. No. 1.4e-121;
  Matches  271;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASPAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 3
Q86UA5
ID   Q86UA5      PRELIMINARY;      PRT;   276 AA.
AC   Q86UA5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   TPS1 protein (Fragment).
GN   Name=TPS1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC051852; AAH51852.1; -.
```

```
DR   HSSP; P15157; 1LTO.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
FT   NON_TER       1      1
SQ   SEQUENCE   276 AA;  30869 MW;  8CDC5EA037955F9C CRC64;

  Query Match             98.8%;  Score 1494;  DB 2;  Length 276;
  Best Local Similarity   98.5%;  Pred. No. 1.4e-121;
  Matches  271;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db          2 MLSLLLLALPVLASPAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 61

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 121

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 181

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||::||||:|||||||||||||||||||
Db        182 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 241

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        242 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 276


RESULT 4
Q6B051
ID   Q6B051      PRELIMINARY;      PRT;   275 AA.
AC   Q6B051;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Beta tryptase 1,.
GN   Name=TPSB1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pooled tissue;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
```

```
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pooled tissue;
RA   Director MGC Project;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC074975; AAH74975.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   275 AA;  30488 MW;  A8F0B12C58F88C12 CRC64;

  Query Match             93.5%;  Score 1414;  DB 2;  Length 275;
  Best Local Similarity   92.7%;  Pred. No. 1.3e-114;
  Matches  255;  Conservative    7;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              |||||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLSLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 5
TRB1_HUMAN
ID   TRB1_HUMAN     STANDARD;      PRT;   275 AA.
AC   Q15661; Q15663; Q9H2Y4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Tryptase beta-1 precursor (EC 3.4.21.59) (Tryptase 1) (Tryptase I).
GN   Name=TPSB1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90251647; PubMed=2187193;
RA   Vanderslice P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
RA   Caughey G.H.;
RT   "Human mast cell tryptase: multiple cDNAs and genes reveal a multigene
RT   serine protease family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99121069; PubMed=9920877; DOI=10.1074/jbc.274.6.3355;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [3]
RP   SEQUENCE OF 54-275 FROM N.A. (ISOFORM 2).
RA   Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA   Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA   Hunt J.E.;
RT   "Molecular cloning and characterization of novel human tryptase cDNAs
RT   and splicing variants.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q15661-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q15661-2; Sequence=VSP_005375;
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Tryptase
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M33494; AAC83172.1; -.
DR   EMBL; M33491; AAA36778.1; -.
DR   EMBL; AF099144; AAD17860.1; -.
DR   EMBL; AF206667; AAG35697.1; -.
DR   PIR; A35863; A35863.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.015; -.
DR   MEROPS; S01.242; -.
DR   Genew; HGNC:12019; TPSB1.
DR   MIM; 191081; -.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Alternative splicing; Glycoprotein; Hydrolase; Polymorphism;
KW   Serine protease; Signal; Zymogen.
FT   SIGNAL        1     18       Potential.
FT   PROPEP       19     30       Activation peptide (By similarity).
FT   CHAIN        31    275       Tryptase beta-1.
FT   ACT_SITE     74     74       Charge relay system.
FT   ACT_SITE    121    121       Charge relay system.
FT   ACT_SITE    224    224       Charge relay system.
FT   DISULFID     59     75       By similarity.
FT   DISULFID    155    230       By similarity.
FT   DISULFID    188    211       By similarity.
FT   DISULFID    220    248       By similarity.
```

```
FT   CARBOHYD    132    132       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC     79     87       Missing (in isoform 2).
FT                                /FTId=VSP_005375.
FT   VARIANT      18     18       A -> V (in dbSNP:1800984).
FT                                /FTId=VAR_014557.
FT   VARIANT      23     23       G -> V (in dbSNP:1800986).
FT                                /FTId=VAR_014558.
FT   VARIANT      85     85       A -> T (in dbSNP:2234902).
FT                                /FTId=VAR_014559.
FT   VARIANT     132    132       N -> K (in dbSNP:1800991).
FT                                /FTId=VAR_016102.
FT   VARIANT     141    141       T -> A (in dbSNP:1800992).
FT                                /FTId=VAR_014560.
FT   VARIANT     162    162       D -> N (in dbSNP:2234641).
FT                                /FTId=VAR_014561.
FT   VARIANT     170    170       P -> S (in dbSNP:2234904).
FT                                /FTId=VAR_014562.
FT   VARIANT     215    215       T -> S (in dbSNP:2234905).
FT                                /FTId=VAR_014563.
FT   VARIANT     216    216       R -> Q (in dbSNP:2234906).
FT                                /FTId=VAR_014564.
SQ   SEQUENCE   275 AA;  30515 MW;  ADC48FDC51F37112 CRC64;

  Query Match              93.3%;  Score 1411;  DB 1;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 2.3e-114;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
                ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
                |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
                ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
                ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
                |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 6
TRB2_HUMAN
ID   TRB2_HUMAN     STANDARD;      PRT;   275 AA.
AC   P20231; O95827; Q15664; Q9UQI6; Q9UQI7;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Tryptase beta-2 precursor (EC 3.4.21.59) (Tryptase 2) (Tryptase II).
GN   Name=TPSB2; Synonyms=TPS2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (VARIANT BETA-2).
RC   TISSUE=Lung;
RX   MEDLINE=90369005; PubMed=2203827;
RA   Miller J.S., Moxley G., Schwartz L.B.;
RT   "Cloning and characterization of a second complementary DNA for human
RT   tryptase.";
RL   J. Clin. Invest. 86:864-870(1990).
RN   [2]
RP   SEQUENCE FROM N.A. (VARIANT BETA-2).
RX   MEDLINE=93166209; PubMed=8434231;
```

```
RA   Blom T., Hellman L.;
RT   "Characterization of a tryptase mRNA expressed in the human basophil
RT   cell line KU812.";
RL   Scand. J. Immunol. 37:203-208(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (VARIANTS BETA-2 AND BETA-3).
RX   MEDLINE=90251647; PubMed=2187193;
RA   Vanderslice P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
RA   Caughey G.H.;
RT   "Human mast cell tryptase: multiple cDNAs and genes reveal a multigene
RT   serine protease family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).
RN   [4]
RP   SEQUENCE FROM N.A. (VARIANTS BETA-2 AND BETA-3).
RX   MEDLINE=99121069; PubMed=9920877; DOI=10.1074/jbc.274.6.3355;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [5]
RP   REVISIONS.
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX   MEDLINE=98180625; PubMed=9521329; DOI=10.1038/32703;
RA   Pereira P.J.B., Bergner A., Macedo-Ribeiro S., Huber R.,
RA   Matschiner G., Fritz H., Sommerhoff C.P., Bode W.;
RT   "Human beta-tryptase is a ring-like tetramer with active sites facing
RT   a central pore.";
RL   Nature 392:306-311(1998).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS).
RX   MEDLINE=99432168; PubMed=10500112; DOI=10.1073/pnas.96.20.10984;
RA   Sommerhoff C.P., Bode W., Pereira P.J.B., Stubbs M.T.,
RA   Stuerzebecher J., Piechottka G.P., Matschiner G., Bergner A.;
RT   "The structure of the human betaII-tryptase tetramer: fo(u)r better or
RT   worse.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10984-10991(1999).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- POLYMORPHISM: There are two alleles; beta-II and beta-III which
CC       differ by 3 residues.
```

```
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Tryptase
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M37488; AAA51843.1; -.
DR   EMBL; M33492; AAA36779.1; -.
DR   EMBL; M33493; AAA36780.1; -.
DR   EMBL; S55551; AAD13876.1; -.
DR   EMBL; AF099143; AAD17859.2; -.
DR   EMBL; AF099145; AAD17857.1; -.
DR   EMBL; AF099146; AAD17858.1; -.
DR   EMBL; BC029356; AAH29356.1; -.
DR   PIR; B35863; B35863.
DR   PDB; 1A0L; X-ray; A/B/C/D=31-274.
DR   PDB; 1AAO; Model; @=31-275.
DR   MEROPS; S01.242; -.
DR   Genew; HGNC:14120; TPSB2.
DR   H-InvDB; HIX0012676; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   3D-structure; Glycoprotein; Hydrolase; Polymorphism; Serine protease;
KW   Signal; Zymogen.
FT   SIGNAL        1     18       Potential.
FT   PROPEP       19     30       Activation peptide.
FT   CHAIN        31    275       Tryptase beta-2.
FT   ACT_SITE     74     74       Charge relay system.
FT   ACT_SITE    121    121       Charge relay system.
FT   ACT_SITE    224    224       Charge relay system.
FT   DISULFID     59     75
FT   DISULFID    155    230
FT   DISULFID    188    211
FT   DISULFID    220    248
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT      51     53       HGP -> RDR (in beta-III).
FT                                /FTId=VAR_012104.
FT   CONFLICT    132    132       K -> N (in Ref. 3).
FT   STRAND       32     32
FT   STRAND       35     36
FT   TURN         39     40
FT   TURN         43     44
FT   STRAND       45     51
FT   STRAND       55     65
FT   TURN         66     67
FT   STRAND       68     71
FT   HELIX        73     75
FT   TURN         76     76
FT   STRAND       79     79
FT   HELIX        83     85
FT   STRAND       86     90
FT   STRAND       94     94
FT   TURN         95     98
FT   STRAND      102    102
FT   STRAND      104    109
FT   TURN        111    112
FT   HELIX       116    118
FT   STRAND      123    127
```

```
FT   STRAND      141    141
FT   TURN        145    146
FT   TURN        151    152
FT   STRAND      155    159
FT   STRAND      164    164
FT   TURN        165    166
FT   STRAND      167    167
FT   TURN        171    172
FT   STRAND      174    174
FT   STRAND      176    180
FT   STRAND      182    183
FT   HELIX       185    191
FT   TURN        192    195
FT   STRAND      196    196
FT   TURN        200    201
FT   TURN        207    208
FT   STRAND      209    212
FT   STRAND      214    214
FT   TURN        215    216
FT   STRAND      217    218
FT   TURN        221    222
FT   TURN        224    225
FT   STRAND      227    232
FT   TURN        233    234
FT   STRAND      235    244
FT   STRAND      250    250
FT   TURN        251    252
FT   STRAND      253    253
FT   STRAND      255    259
FT   HELIX       260    263
FT   HELIX       264    270
SQ   SEQUENCE   275 AA;  30529 MW;  2B27396C51F5C7A0 CRC64;

  Query Match             92.9%;  Score 1405;  DB 1;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 7.7e-114;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 7
Q6NZY1
ID   Q6NZY1      PRELIMINARY;      PRT;   282 AA.
AC   Q6NZY1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   TPSB2 protein.
GN   Name=TPSB2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Director MGC Project;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC065923; AAH65923.1; -.
DR   HSSP; P20231; 1A0L.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   282 AA;  31296 MW;  9727DC527F9896CA CRC64;

  Query Match             92.9%;  Score 1405;  DB 2;  Length 282;
  Best Local Similarity   92.0%;  Pred. No. 7.9e-114;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          8 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 67

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         68 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 127

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        128 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 187

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        188 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 247

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        248 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 282


RESULT 8
Q6FHB8
ID   Q6FHB8       PRELIMINARY;      PRT;   275 AA.
```

```
AC   Q6FHB8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   TPSB2 protein (Fragment).
GN   Name=TPSB2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Halleck A., Ebert L., Mkoundinya M., Schick M., Eisenstein S.,
RA   Neubert P., Kstrang K., Schatten R., Shen B., Henze S., Mar W.,
RA   Korn B., Zuo D., Hu Y., LaBaer J.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; CR541836; CAG46635.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
FT   NON_TER     275    275
SQ   SEQUENCE   275 AA;  30517 MW;  ABD124C4293868BD CRC64;

  Query Match             92.8%;  Score 1403;  DB 2;  Length 275;
  Best Local Similarity   91.6%;  Pred. No. 1.2e-113;
  Matches  252;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |:|||||||||||||:|||||||| ||||||||||||||||||||| |||||||  | ||
Db         61 GTLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRTIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 9
Q96RZ6
ID   Q96RZ6       PRELIMINARY;      PRT;   275 AA.
AC   Q96RZ6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Tryptase I.
GN   Name=tryptaseC;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797; DOI=10.1093/hmg/10.4.339;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AE006466; AAK61271.1; -.
DR   HSSP; P20231; 1A0L.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   275 AA;  30601 MW;  2F4AAAAE7DDE483E CRC64;

  Query Match             92.7%;  Score 1401;  DB 2;  Length 275;
  Best Local Similarity   91.6%;  Pred. No. 1.7e-113;
  Matches  252;  Conservative    9;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||:|: ||||||||||||||||||||| | |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQAVQRVGIVGGQEAPRSKWPWQVSLRVHDPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              || | ||||||||||||||||||||||| ||||||
Db        241 VVRWGEGCAQPNRPGIYTRVTYYLDWIHXYVPKKP 275


RESULT 10
Q7YS62
ID   Q7YS62      PRELIMINARY;      PRT;   275 AA.
AC   Q7YS62;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Tryptase (EC 3.4.21.59).
GN   Name=mtc1;
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dacre K.J., McAleese S., Pemberton A., McGorum B.C.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Dacre K.J.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ515902; CAD56807.1; -.
DR   HSSP; P20231; 1A0L.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0030019; F:tryptase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   275 AA;  30476 MW;  718F57A4DBEE21ED CRC64;

  Query Match             76.5%;  Score 1156;  DB 2;  Length 275;
  Best Local Similarity   74.4%;  Pred. No. 3.3e-92;
  Matches  203;  Conservative   28;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              | :||:||| :| :  :||||| |||:: ||||||||  ||||||||||    || ||||
Db          1 MPNLLVLALALLVNLGHAAPAPGQALEREGIVGGQEASGSKWPWQVSLRKNTEYWKHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||::|   :|||||||||||:|||||||||: || :| :: ||
Db         61 GSLIHPQWVLTAAHCVGPDIEDFRDIRVQLREQHLYYRDQLLPVSRILPHPYYYTVENGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              |||||||::|||||| |  | ||||||||||| ||||||||||||   ||||||||:|||
Db        121 DIALLELQDPVNISSHVQVVTLPPASETFPPGTPCWVTGWGDVDNGVSLPPPFPLKEVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||:|| :|| ||| |  |||::||:: ||||||| : |||:||||||||||| |||||||
Db        181 PIVENSVCDRKYHTGVSTGDNIRIVQADMLCAGNRRHDSCQGDSGGPLVCKVKGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              ||||   ||||||||||||||||||||: ||||
Db        241 VVSWANSCAQPNRPGIYTRVTYYLDWIYQYVPK 273


RESULT 11
MCT7_MOUSE
ID   MCT7_MOUSE     STANDARD;      PRT;   273 AA.
AC   Q02844;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Mast cell protease 7 precursor (EC 3.4.21.59) (MMCP-7) (Tryptase).
GN   Name=Mcpt7;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2;
RX   MEDLINE=93087489; PubMed=1454796;
RA   McNeil H.P., Reynolds D.S., Schiller V., Ghildyal N., Gurley D.S.,
RA   Austen K.F., Stevens R.L.;
RT   "Isolation, characterization, and transcription of the gene encoding
RT   mouse mast cell protease 7.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:11174-11178(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J;
RX   MEDLINE=96162035; PubMed=8576265; DOI=10.1074/jbc.271.5.2851;
RA   Hunt J.E., Stevens R.L., Austen K.F., Zhang J., Xia Z., Ghildyal N.;
RT   "Natural disruption of the mouse mast cell protease 7 gene in the
RT   C57BL/6 mouse.";
RL   J. Biol. Chem. 271:2851-2855(1996).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
```

```
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=The alternative splicing event is due to a G to A point
CC         mutation at the exon 2/intron 2 splice site and causes loss of
CC         protein expression. The alternatively spliced transcript is only
CC         found in C57BL/6 mouse;
CC       Name=1;
CC         IsoId=Q02844-1; Sequence=Displayed;
CC       Name=2; Synonyms=Truncated;
CC         IsoId=Q02844-2; Sequence=VSP_005378, VSP_005379;
CC   -!- DEVELOPMENTAL STAGE: Is not expressed in mature serosal or mucosal
CC       mast cells and is expressed only transiently at an early stage of
CC       in vitro mast cell differentiation.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Tryptase
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L00654; AAA39993.1; -.
DR   EMBL; L00653; AAA39992.1; -.
DR   EMBL; U42405; AAA97874.1; -.
DR   EMBL; U42406; AAA97875.1; -.
DR   PIR; A47246; A47246.
DR   HSSP; P15157; 1LTO.
DR   MEROPS; S01.026; -.
DR   MGD; MGI:96943; Mcpt7.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Alternative splicing; Glycoprotein; Hydrolase; Serine protease;
KW   Signal; Zymogen.
FT   SIGNAL        1     18       Potential.
FT   PROPEP       19     28       Activation peptide (Potential).
FT   CHAIN        29    273       Mast cell protease 7.
FT   ACT_SITE     72     72       Charge relay system (By similarity).
FT   ACT_SITE    119    119       Charge relay system (By similarity).
FT   ACT_SITE    222    222       Charge relay system (By similarity).
FT   DISULFID     57     73       By similarity.
FT   DISULFID    153    228       By similarity.
FT   DISULFID    186    209       By similarity.
FT   DISULFID    218    246       By similarity.
FT   CARBOHYD     49     49       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    130    130       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC     44     46       VSL -> GCC (in isoform 2).
FT                                /FTId=VSP_005378.
FT   VARSPLIC     47    273       Missing (in isoform 2).
FT                                /FTId=VSP_005379.
SQ   SEQUENCE   273 AA;  30337 MW;  50ECB4765294205E CRC64;

  Query Match             75.0%;  Score 1134;  DB 1;  Length 273;
  Best Local Similarity   74.4%;  Pred. No. 2.7e-90;
  Matches  203;  Conservative   23;  Mismatches   45;  Indels    2;  Gaps    1;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              || |||| ||:|:|  :||| |  |: : ||||||||  :|||||||||  | |||||||
Db          1 MLKLLLLTLPLLSSLVHAAPGP--AMTREGIVGGQEAHGNKWPWQVSLRANDTYWMHFCG 58

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||| |   :|||||:|:||| | |: ||:|| || |||:| ||
Db         59 GSLIHPQWVLTAAHCVGPDVADPNKVRVQLRKQYLYYHDHLMTVSQIITHPDFYIVQDGA 118

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              |||||:|  |||||  || | ||||||||| |  |||||||::||   ||||||||:|:|
Db        119 DIALLKLTNPVNISDYVHPVPLPPASETFPSGTLCWVTGWGNIDNGVNLPPPFPLKEVQV 178

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||:|||:|| ||| |  |||:| |:|||||||||   |||:|||||||||||  ||||||
Db        179 PIIENHLCDLKYHKGLITGDNVHIVRDDMLCAGNEGHDSCQGDSGGPLVCKVEDTWLQAG 238

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              |||| ||||||||||||||||||||||||||||
Db        239 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPK 271


RESULT 12
MCT6_MOUSE
ID   MCT6_MOUSE     STANDARD;      PRT;   276 AA.
AC   P21845; Q61962;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Mast cell protease 6 precursor (EC 3.4.21.59) (MMCP-6) (Tryptase).
GN   Name=Mcpt6;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91139682; PubMed=1995638;
RA   Reynolds D.S., Gurley D.S., Austen K.F., Serafin W.E.;
RT   "Cloning of the cDNA and gene of mouse mast cell protease-6.
RT   Transcription by progenitor mast cells and mast cells of the
RT   connective tissue subclass.";
RL   J. Biol. Chem. 266:3847-3853(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Leaden X A1;
RX   MEDLINE=94023807; PubMed=8210998;
RA   Huang R., Abrink M., Gobl A.E., Nilsson G., Aveskogh M., Larsson L.G.,
RA   Nilsson K., Hellman L.;
RT   "Expression of a mast cell tryptase in the human monocytic cell lines
RT   U-937 and Mono Mac 6.";
RL   Scand. J. Immunol. 38:359-367(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   STRAIN=Leaden X A1;
RX   MEDLINE=95048582; PubMed=7959952;
RA   Huang R., Hellman L.T.;
RT   "Genes for mast-cell serine protease and their molecular evolution.";
RL   Immunogenetics 40:397-414(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RC   STRAIN=FVB/N; TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

```
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 32-54.
RX   MEDLINE=90222202; PubMed=2326280;
RA   Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F.,
RA   Serafin W.E.;
RT   "Different mouse mast cell populations express various combinations of
RT   at least six distinct mast cell serine proteases.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3230-3234(1990).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P21845-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P21845-2; Sequence=VSP_005376, VSP_005377;
CC         Note=Probably non functional;
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Tryptase
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M57626; AAA39988.1; -.
DR   EMBL; M57625; AAA39987.1; -.
DR   EMBL; L31853; AAA39725.1; -.
DR   EMBL; X78542; CAA55288.1; -.
DR   EMBL; BC024374; AAH24374.1; -.
DR   PIR; A38654; A38654.
DR   PIR; I48685; I48685.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.025; -.
DR   MGD; MGI:96942; Mcpt6.
DR   GO; GO:0008201; F:heparin binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0030019; F:tryptase activity; IDA.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Alternative splicing; Direct protein sequencing; Glycoprotein;
KW   Hydrolase; Serine protease; Signal; Zymogen.
FT   SIGNAL        1     21
FT   PROPEP       22     31       Activation peptide.
FT   CHAIN        32    276       Mast cell protease 6.
FT   ACT_SITE     75     75       Charge relay system (By similarity).
FT   ACT_SITE    122    122       Charge relay system (By similarity).
FT   ACT_SITE    225    225       Charge relay system (By similarity).
FT   DISULFID     60     76       By similarity.
FT   DISULFID    156    231       By similarity.
FT   DISULFID    189    212       By similarity.
FT   DISULFID    221    249       By similarity.
FT   CARBOHYD    133    133       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC    223    230       GDSGGPLV -> PFCIGDDI (in isoform Short).
FT                                /FTId=VSP_005376.
FT   VARSPLIC    231    276       Missing (in isoform Short).
```

```
FT                                /FTId=VSP_005377.
SQ   SEQUENCE   276 AA;  30927 MW;  525B2C9A04A72200 CRC64;

  Query Match             74.9%;  Score 1133;  DB 1;  Length 276;
  Best Local Similarity   73.3%;  Pred. No. 3.3e-90;
  Matches  198;  Conservative   28;  Mismatches   42;  Indels    2;  Gaps    1;

Qy          4 LLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSL 63
              ||| || :|||  |:|| |  | |: ||||| ||  |||||||||| :  ||:|||||||
Db          7 LLLWALSLLASLVYSAPRP--ANQRVGIVGGHEASESKWPWQVSLRFKLNYWIHFCGGSL 64

Qy         64 IHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIA 123
              ||||||||||||:|| :|     |||||||:||| |||| ::||:||| :|  : |||:|
Db         65 IHPQWVLTAAHCVGPHIKSPQLFRVQLREQYLYYGDQLLSLNRIVVHPHYYTAEGGADVA 124

Qy        124 LLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIM 183
              ||||| |||:|: :| : |||||||||||  ||||||||:|||||||||:|||||||||:
Db        125 LLELEVPVNVSTHIHPISLPPASETFPPGTSCWVTGWGDIDNDEPLPPPYPLKQVKVPIV 184

Qy        184 ENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVS 243
              ||  :|| ||| | |||||  |: | ||||||::||||:||||||||||| ||||||||||
Db        185 ENSLCDRKYHTGLYTGDDFPIVHDGMLCAGNTRRDSCQGDSGGPLVCKVKGTWLQAGVVS 244

Qy        244 WDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              | |||||||:||||||||||||||| |||:
Db        245 WGEGCAQPNKPGIYTRVTYYLDWIHRYVPE 274


RESULT 13
TRYT_CANFA
ID   TRYT_CANFA     STANDARD;      PRT;   275 AA.
AC   P15944;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Tryptase precursor (EC 3.4.21.59).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89352460; PubMed=2504277;
RA   Vanderslice P., Craik C.S., Nadel J.A., Caughey G.H.;
RT   "Molecular cloning of dog mast cell tryptase and a related protease:
RT   structural evidence of a unique mode of serine protease activation.";
RL   Biochemistry 28:4148-4155(1989).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Tryptase
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M24664; AAA30854.1; -.
DR   PIR; A32410; A32410.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.118; -.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
```

```
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Signal; Zymogen.
FT   SIGNAL        1     20       Potential.
FT   PROPEP       21     30       Activation peptide (By similarity).
FT   CHAIN        31    275       Tryptase.
FT   ACT_SITE     74     74       Charge relay system (By similarity).
FT   ACT_SITE    121    121       Charge relay system (By similarity).
FT   ACT_SITE    224    224       Charge relay system (By similarity).
FT   DISULFID     59     75       By similarity.
FT   DISULFID    155    230       By similarity.
FT   DISULFID    188    211       By similarity.
FT   DISULFID    220    248       By similarity.
FT   CARBOHYD    132    132       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   275 AA;  30088 MW;  C3B869251F248D5B CRC64;

  Query Match             74.8%;  Score 1131;  DB 1;  Length 275;
  Best Local Similarity   72.7%;  Pred. No. 4.9e-90;
  Matches  200;  Conservative   30;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              | | |:||| :| |    :||| ||||: |||||:||| ||||||||||:: :|| | ||
Db          1 MPSPLVLALALLGSLVPVSPAPGQALQRVGIVGGREAPGSKWPWQVSLRLKGQYWRHICG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:||:|     :||||||||||||| ||||:||::|| :|  : ||
Db         61 GSLIHPQWVLTAAHCVGPNVVCPEEIRVQLREQHLYYQDHLLPVNRIVMHPNYYTPENGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||:|||:|: |  | |||| :||| | |||||||||| :  |||||||||||||
Db        121 DIALLELEDPVNVSAHVQPVTLPPALQTFPTGTPCWVTGWGDVHSGTPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||:|| :|| :||||  ||| |||:|:||||||||: |||:|||||||||:| | |||||
Db        181 PIVENSMCDVQYHLGLSTGDGVRIVREDMLCAGNSKSDSCQGDSGGPLVCRVRGVWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||| ||||||| ||||:|
Db        241 VVSWGEGCAQPNRPGIYTRVAYYLDWIHQYVPKEP 275

RESULT 14
Q921N4
ID   Q921N4       PRELIMINARY;      PRT;   273 AA.
AC   Q921N4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mast cell protease 7.
GN   Name=Mcpt7;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

```
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC011328; AAH11328.1; -.
DR   HSSP; P15157; 1LTO.
DR   MGD; MGI:96943; Mcpt7.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   273 AA;  30332 MW;  51417476514035BE CRC64;

  Query Match             74.3%;  Score 1123;  DB 2;  Length 273;
  Best Local Similarity   74.0%;  Pred. No. 2.4e-89;
  Matches  202;  Conservative   22;  Mismatches   47;  Indels    2;  Gaps    1;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              || |||| ||:|:|   ||| |  |: : ||||||||  :|||||||||  | |||||||
Db          1 MLKLLLLTLPLLSSLVLAAPGP--AMTREGIVGGQEAHGNKWPWQVSLRANDTYWMHFCG 58

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||| |   :|||||:|:||| | |: ||:|| || |||:| ||
Db         59 GSLIHPQWVLTAAHCVGPDVADPNKVRVQLRKQYLYYHDHLMTVSQIITHPDFYIVQDGA 118

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              |||||:|  |||||  || | |||||||||| |  |||||||::||   ||||||||:|:|
Db        119 DIALLKLTNPVNISDYVHPVPLPPASETFPSGTLCWVTGWGNIDNGVNLPPPFPLKEVQV 178

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||:|||:|| ||| |  |||:| |:|||||||||   |||:|||||||||||  ||||||
Db        179 PIIENHLCDLKYHKGLITGDNVHIVRDDMLCAGNEGHDSCQGDSGGPLVCKVEDTWLQAG 238

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPK 273
              |||| ||||||||||||||||||||||| ||||
Db        239 VVSWGEGCAQPNRPGIYTRVTYYLDWIHRYVPK 271

RESULT 15
Q96RZ7
ID   Q96RZ7       PRELIMINARY;      PRT;   233 AA.
AC   Q96RZ7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mast cell tryptase beta III.
GN   Name=tryptaseB;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797; DOI=10.1093/hmg/10.4.339;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
```

```
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AE006466; AAK61270.1; -.
DR   HSSP; P20231; 1A0L.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   233 AA;  25874 MW;  246A96C71B2CE6D0 CRC64;

  Query Match             74.1%;  Score 1120;  DB 2;  Length 233;
  Best Local Similarity   92.3%;  Pred. No. 3.7e-89;
  Matches  204;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||||||||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCK 221
              ||||||||||||||||||||||||:|||||||||::||||:
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQ 221


Search completed: August 27, 2005, 23:29:07
Job time : 133 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         August 27, 2005, 23:39:06 ; Search time 615 Seconds
                                           (without alignments)
                                           2925.772 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters:      14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_25082005_165827_1602/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09598982_@CGN_1_1_723_@runat_25082005_165827_1602
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
                19:  /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
                20:  /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
                21:  /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
                22:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                23:  /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
                24:  /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
                25:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                26:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
```

```
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                      %
Result              Query
  No.    Score     Match Length DB  ID                          Description
---------------------------------------------------------------------------------
    1     1512     100.0   1154 17  US-10-352-684A-45           Sequence 45, Appl
    2     1512     100.0   1158 14  US-10-116-802-240           Sequence 240, App
    3     1512     100.0   1158 21  US-10-956-157-1042          Sequence 1042, Ap
    4     1416      93.7   1334 20  US-10-723-860-6799          Sequence 6799, Ap
    5     1414      93.5    828 18  US-10-287-226-89            Sequence 89, Appl
    6     1411      93.3    828 18  US-10-287-226-91            Sequence 91, Appl
    7     1411      93.3   1193 21  US-10-956-157-1043          Sequence 1043, Ap
    8     1405      92.9   1143 17  US-10-352-684A-47           Sequence 47, Appl
    9     1405      92.9   1145 18  US-10-287-226-93            Sequence 93, Appl
   10     1405      92.9   1148 21  US-10-956-157-2444          Sequence 2444, Ap
   11     1401      92.7   2662 18  US-10-275-505-27            Sequence 27, Appl
   12     1394      92.2   1081 9  US-09-954-456-2126          Sequence 2126, Ap
   13     1394      92.2   1081 10  US-09-960-706-680           Sequence 680, App
   14     1394      92.2   1081 10  US-09-873-319-427           Sequence 427, App
   15     1394      92.2   1081 10  US-09-873-367C-155          Sequence 155, App
   16     1394      92.2   1081 10  US-09-873-367C-714          Sequence 714, App
   17     1394      92.2   1081 21  US-10-843-641A-155          Sequence 155, App
   18     1394      92.2   1081 21  US-10-843-641A-714          Sequence 714, App
   19     1394      92.2   1081 21  US-10-843-641A-5153         Sequence 5153, Ap
   20   1356.5      89.7    858 18  US-10-287-226-87            Sequence 87, Appl
   21   1246.5      82.4   5456 20  US-10-723-860-2689          Sequence 2689, Ap
   22   1246.5      82.4   5456 22  US-10-756-149-2484          Sequence 2484, Ap
   23     1134      75.0   1187 18  US-10-275-505-22            Sequence 22, Appl
   24     1108      73.3    729 21  US-10-480-988-55            Sequence 55, Appl
   25     1100      72.8    729 21  US-10-956-157-1789          Sequence 1789, Ap
   26      953      63.0    600 21  US-10-956-157-7024          Sequence 7024, Ap
   27    687.5      45.5    846 10  US-09-813-432-11            Sequence 11, Appl
   28    687.5      45.5    846 17  US-10-174-364-11            Sequence 11, Appl
   29    687.5      45.5    846 18  US-10-246-583-11            Sequence 11, Appl
   30    687.5      45.5    846 19  US-10-689-832-11            Sequence 11, Appl
   31      643      42.5    873 14  US-10-117-323-2             Sequence 2, Appli
   32    632.5      41.8   1199 21  US-10-956-157-1816          Sequence 1816, Ap
   33      616      40.7    825 14  US-10-117-323-1             Sequence 1, Appli
   34    606.5      40.1    948 16  US-10-131-409-15            Sequence 15, Appl
   35    606.5      40.1    948 17  US-10-139-854-15            Sequence 15, Appl
   36    606.5      40.1    948 17  US-10-150-813-15            Sequence 15, Appl
   37    606.5      40.1    948 17  US-10-150-811-15            Sequence 15, Appl
   38    604.5      40.0   1958 17  US-10-311-035-29            Sequence 29, Appl
   39    603.5      39.9   1122 9  US-09-900-754-1             Sequence 1, Appli
   40      601      39.7    858 16  US-10-131-409-21            Sequence 21, Appl
   41      601      39.7    858 17  US-10-139-854-21            Sequence 21, Appl
   42      601      39.7    858 17  US-10-150-813-21            Sequence 21, Appl
   43      601      39.7    858 17  US-10-150-811-21            Sequence 21, Appl
   44      588      38.9    383 10  US-09-918-995-6378          Sequence 6378, Ap
   45    587.5      38.9    870 16  US-10-131-409-19            Sequence 19, Appl


                                  ALIGNMENTS

RESULT 1
US-10-352-684A-45
; Sequence 45, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
```

```
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 1154
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (18)...(845)
US-10-352-684A-45

Alignment Scores:
Pred. No.:               5.51e-171     Length:        1154
Score:                   1512.00       Matches:       275
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      17            Gaps:          0

US-09-598-982C-52 (1-275) x US-10-352-684A-45 (1-1154)

Qy            1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy           21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy           41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy           61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy           81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy          101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy          121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy          141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy          161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy          181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy          201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy          221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy          241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy          261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
                |||||||||||||||||||||||||||||||||||||||||||||
Db          798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 2
US-10-116-802-240
; Sequence 240, Application US/10116802
; Publication No. US20030065157A1
; GENERAL INFORMATION:
;  APPLICANT: Amy Lasek
;  TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
;  FILE REFERENCE: PA-0045 US
;  CURRENT APPLICATION NUMBER: US/10/116,802
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 60/281,593
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 519
;  SOFTWARE: PERL Program
; SEQ ID NO 240
;   LENGTH: 1158
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 378633.40
US-10-116-802-240

Alignment Scores:
Pred. No.:               5.54e-171     Length:        1158
Score:                   1512.00       Matches:       275
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      14            Gaps:          0

US-09-598-982C-52 (1-275) x US-10-116-802-240 (1-1158)

Qy            1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy           21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy           41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy           61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 3
US-10-956-157-1042
; Sequence 1042, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1042
;   LENGTH: 1158
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-1042

Alignment Scores:
Pred. No.:              5.54e-171      Length:         1158
Score:                  1512.00        Matches:        275
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     21             Gaps:           0

US-09-598-982C-52 (1-275) x US-10-956-157-1042 (1-1158)
```

```
Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 4
US-10-723-860-6799
; Sequence 6799, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 6799
;   LENGTH: 1334
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-6799

Alignment Scores:
Pred. No.:              2.23e-159      Length:         1334
Score:                  1416.00        Matches:        255
Percent Similarity:     95.64%         Conservative:   8
Best Local Similarity:  92.73%         Mismatches:     12
Query Match:            93.65%         Indels:         0
DB:                     20             Gaps:           0

US-09-598-982C-52 (1-275) x US-10-723-860-6799 (1-1334)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 112

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db        113 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 172

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              |||||||||||||||||||||||||||||||||      |||||||||||||||||||||
Db        173 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGGCCCATACTGGATGCACTTCTGCGGG 232

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        233 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 292

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        293 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 352

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        353 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 412

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        413 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 472

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 532

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        533 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 592

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 652

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        653 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 712

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 772

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 832

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        833 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 877

RESULT 5
```

```
US-10-287-226-89
; Sequence 89, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 89
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-89

Alignment Scores:
Pred. No.:              2.05e-159     Length:        828
Score:                  1414.00       Matches:       255
Percent Similarity:     95.64%        Conservative:  8
Best Local Similarity:  92.73%        Mismatches:    12
Query Match:            93.52%        Indels:        0
DB:                     18            Gaps:          0

US-09-598-982C-52 (1-275) x US-10-287-226-89 (1-828)

Qy            1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
                ||||||||||||||||||||||||||||||||||||||||||   ||||||   ||||||
Db            1 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGTGGCCCCT 60

Qy           21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                ||||||   |||||||||||||||   |||||||||||||||||||||||||||||||||
Db           61 GCCCCAGGCCAGGCCCTGCAGCAAACGGGCATTGTTGGGGGGCAGGAGGCCCCCAGGAGC 120

Qy           41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
                |||||||||||||||||||||||||||||||||      |||||||||||||||||||||
Db          121 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGGCCCATACTGGATGCACTTCTGCGGG 180

Qy           61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
                |||||||||||||||||||||||||||||||||||||||||||||:::   ||||||:::
Db          181 GGCTCCCTCATCCACCCCCAGTGGGTGCTAACCGCGGCGCACTGCGTGGAACCGGACATC 240

Qy           81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db          241 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 300

Qy          101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACCGGGGCG 360

Qy          121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
                |||||||||||||||||||||||||||||||||||||||||||||   :::|||||||||
Db          361 GACATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCACATCCACACGGTC 420

Qy          141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 ACGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 480

Qy          161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
                |||||||||||||||:::|||   ||||||||||||||||||||||||||||||||||||
Db          481 GGCGACGTGGACAATAATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 540

Qy          181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 600

Qy          201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
                ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db          601 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 660

Qy          221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
                :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 720

Qy          241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db          721 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 780

Qy          261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
                |||||||||||||||||||||||||||||||||||||||||||||
Db          781 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 825

RESULT 6
US-10-287-226-91
; Sequence 91, Application US/10287226
```

```
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 91
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
```

```
US-10-287-226-91

Alignment Scores:
Pred. No.:              4.69e-159      Length:        828
Score:                  1411.00        Matches:       254
Percent Similarity:     95.27%         Conservative:  8
Best Local Similarity:  92.36%         Mismatches:    13
Query Match:            93.32%         Indels:        0
DB:                     18             Gaps:          0

US-09-598-982C-52 (1-275) x US-10-287-226-91 (1-828)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 60

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         61 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 120

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        121 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 180

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        181 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 240

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 300

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        301 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 360

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        361 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 420

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 480

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        481 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 540

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 600

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        601 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 660

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 720

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 780

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 825

RESULT 7
US-10-956-157-1043
; Sequence 1043, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1043
;   LENGTH: 1193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-1043

Alignment Scores:
Pred. No.:              7.65e-159      Length:        1193
Score:                  1411.00        Matches:       254
Percent Similarity:     95.27%         Conservative:  8
Best Local Similarity:  92.36%         Mismatches:    13
Query Match:            93.32%         Indels:        0
DB:                     21             Gaps:          0

US-09-598-982C-52 (1-275) x US-10-956-157-1043 (1-1193)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 112

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db        113 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 172

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        173 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 232

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        233 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 292

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        293 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 352

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        353 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 412

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        413 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 472

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 532

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        533 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 592

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 652

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        653 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 712

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 772
```

```
Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 832

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        833 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 877

RESULT 8
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
;   LENGTH: 1143
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6)...(833)
US-10-352-684A-47

Alignment Scores:
Pred. No.:               3.79e-158     Length:        1143
Score:                   1405.00       Matches:       253
Percent Similarity:      94.91%        Conservative:  8
Best Local Similarity:   92.00%        Mismatches:    14
Query Match:             92.92%        Indels:        0
DB:                      17            Gaps:          0

US-09-598-982C-52 (1-275) x US-10-352-684A-47 (1-1143)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 65

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         66 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 125

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        126 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 185

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        186 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 245

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        246 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 305

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        306 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 365

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        366 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 425

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 485

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        486 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 545

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 605

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        606 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 665

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 725

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 785

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        786 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 830

RESULT 9
US-10-287-226-93
; Sequence 93, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
```

```
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 93
;   LENGTH: 1145
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(832)
US-10-287-226-93

Alignment Scores:
Pred. No.:              3.8e-158        Length:         1145
Score:                  1405.00         Matches:        253
Percent Similarity:     94.91%          Conservative:   8
Best Local Similarity:  92.00%          Mismatches:     14
Query Match:            92.92%          Indels:         0
DB:                     18              Gaps:           0

US-09-598-982C-52 (1-275) x US-10-287-226-93 (1-1145)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 67

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         68 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 127
```

```
Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        128 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 187

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        188 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 247

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        248 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 307

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        308 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 367

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        368 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 427

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 487

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        488 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 547

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 607

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        608 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 667

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 727

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        728 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 787

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        788 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 832

RESULT 10
US-10-956-157-2444
; Sequence 2444, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2444
;   LENGTH: 1148
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-2444

Alignment Scores:
Pred. No.:              3.81e-158       Length:         1148
Score:                  1405.00         Matches:        253
Percent Similarity:     94.91%          Conservative:   8
```

```
Best Local Similarity:  92.00%        Mismatches:   14
Query Match:            92.92%        Indels:        0
DB:                     21            Gaps:          0

US-09-598-982C-52 (1-275) x US-10-956-157-2444 (1-1148)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 67

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         68 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 127

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        128 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 187

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        188 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 247

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        248 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 307

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        308 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 367

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        368 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 427

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 487

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        488 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 547

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 607

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        608 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 667

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 727

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        728 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 787

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        788 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 832

RESULT 11
US-10-275-505-27
; Sequence 27, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT:  INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
```

```
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 27
;   LENGTH: 2662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7472460CB1
US-10-275-505-27

Alignment Scores:
Pred. No.:              3.54e-157     Length:        2662
Score:                  1401.00       Matches:       253
Percent Similarity:     94.91%        Conservative:  8
Best Local Similarity:  92.00%        Mismatches:    12
Query Match:            92.66%        Indels:        2
DB:                     18            Gaps:          1

US-09-598-982C-52 (1-275) x US-10-275-505-27 (1-2662)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 676

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                       ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db        677 GGC------CAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 730

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 790

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        791 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 850

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        851 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 910

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        911 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 970

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        971 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 1030

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1031 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 1090

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||   |||||||||
Db       1091 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCATGTGAAGGTC 1150

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 1210

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db       1211 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 1270

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 1330

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||   |||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1331 GTGGTCAGATGGGGAGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 1390

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1391 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 1435

RESULT 12
US-09-954-456-2126
; Sequence 2126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-2126

Alignment Scores:
Pred. No.:              7.36e-157      Length:        1081
Score:                  1394.00        Matches:       249
Percent Similarity:     95.88%         Conservative:  7
Best Local Similarity:  93.26%         Mismatches:    11
Query Match:            92.20%         Indels:        0
DB:                     9              Gaps:          0
```

```
US-09-598-982C-52 (1-275) x US-09-954-456-2126 (1-1081)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db          2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121

Qy         49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy         69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
              |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy         89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy        109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
              ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db        302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy        129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
              |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db        362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy        149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAAGCCG 802

RESULT 13
US-09-960-706-680
; Sequence 680, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasi
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
```

```
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
US-09-960-706-680

Alignment Scores:
Pred. No.:              7.36e-157      Length:        1081
Score:                  1394.00        Matches:       249
Percent Similarity:     95.88%         Conservative:  7
Best Local Similarity:  93.26%         Mismatches:    11
Query Match:            92.20%         Indels:        0
DB:                     10             Gaps:          0

US-09-598-982C-52 (1-275) x US-09-960-706-680 (1-1081)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db          2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121

Qy         49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy         69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
              |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy         89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy        109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
              ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db        302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy        129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
              |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db        362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy        149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAGCCG 802

RESULT 14
US-09-873-319-427
; Sequence 427, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 427
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
US-09-873-319-427

Alignment Scores:
Pred. No.:              7.36e-157      Length:        1081
Score:                  1394.00        Matches:       249
Percent Similarity:     95.88%         Conservative:  7
Best Local Similarity:  93.26%         Mismatches:    11
Query Match:            92.20%         Indels:        0
DB:                     10             Gaps:          0

US-09-598-982C-52 (1-275) x US-09-873-319-427 (1-1081)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db          2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121

Qy         49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy         69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
              |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy         89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy        109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
              ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db        302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy        129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
              |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db        362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy        149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAAGCCG 802

RESULT 15
US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-155

Alignment Scores:
Pred. No.:              7.36e-157     Length:         1081
Score:                  1394.00       Matches:        249
Percent Similarity:     95.88%        Conservative:   7
Best Local Similarity:  93.26%        Mismatches:     11
Query Match:            92.20%        Indels:         0
DB:                     10            Gaps:           0

US-09-598-982C-52 (1-275) x US-09-873-367C-155 (1-1081)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db          2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121

Qy         49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy         69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
              |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy         89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy        109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
              ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db        302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy        129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
              |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db        362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy        149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
Db        422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAAGCCG 802

Search completed: August 28, 2005, 01:30:05
Job time : 622 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         August 27, 2005, 23:30:16 ; Search time 186 Seconds
                                          (without alignments)
                                          2419.226 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:      2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_25082005_165826_1568/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_105_@runat_25082005_165826_1568 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                    Description
 ---------------------------------------------------------------------------
    1    1512   100.0   1154  2  US-09-016-366A-16       Sequence 16, Appl
    2    1512   100.0   1154  2  US-08-978-404B-11       Sequence 11, Appl
    3    1402    92.7   1137  2  US-09-016-366A-18       Sequence 18, Appl
    4    1402    92.7   1137  2  US-08-978-404B-13       Sequence 13, Appl
    5    1400    92.6   1128  2  US-09-016-366A-20       Sequence 20, Appl
    6    1400    92.6   1128  2  US-08-978-404B-15       Sequence 15, Appl
    7    1394    92.2   1081  2  US-09-016-366A-22       Sequence 22, Appl
    8    1394    92.2   1081  2  US-08-978-404B-17       Sequence 17, Appl
    9    1394    92.2   1081  4  US-09-917-254-50       Sequence 50, Appl
   10    1284    84.9    735  3  US-09-079-970A-1        Sequence 1, Appli
   11    1284    84.9    771  3  US-09-079-970A-4        Sequence 4, Appli
   12    1134    75.0   1031  2  US-08-978-404B-1        Sequence 1, Appli
   13    1133    74.9   1108  2  US-09-016-366A-14       Sequence 14, Appl
   14    1133    74.9   1108  2  US-08-978-404B-20       Sequence 20, Appl
   15  1118.5    74.0   2259  2  US-08-845-998-3        Sequence 3, Appli
   16  1118.5    74.0   2259  3  US-09-206-537-3        Sequence 3, Appli
   17  1118.5    74.0   2259  3  US-09-430-854-3        Sequence 3, Appli
   18  1109.5    73.4   1219  2  US-08-978-404B-7        Sequence 7, Appli
   19    1105    73.1   1097  2  US-08-978-404B-4        Sequence 4, Appli
   20    1105    73.1   1103  2  US-09-016-366A-24       Sequence 24, Appl
   21  1080.5    71.5   2218  2  US-08-845-998-5        Sequence 5, Appli
   22  1080.5    71.5   2218  3  US-09-206-537-5        Sequence 5, Appli
   23  1080.5    71.5   2218  3  US-09-430-854-5        Sequence 5, Appli
   24     986    65.2   3757  2  US-09-016-366A-13       Sequence 13, Appl
   25     986    65.2   3757  2  US-08-978-404B-19       Sequence 19, Appl
   26   892.5    59.0   2397  2  US-08-978-404B-2        Sequence 2, Appli
   27   686.5    45.4   1095  2  US-08-978-404B-9        Sequence 9, Appli
   28   572.5    37.9   1110  3  US-09-386-653A-1        Sequence 1, Appli
   29   570.5    37.7   1212  4  US-09-620-312D-431      Sequence 431, App
   30   560.5    37.1    980  4  US-09-023-942A-30       Sequence 30, Appl
   31     555    36.7   1130  3  US-09-386-653A-8        Sequence 8, Appli
   32   553.5    36.6    933  4  US-09-023-942A-29       Sequence 29, Appl
   33   553.5    36.6   1613  4  US-09-387-375-1        Sequence 1, Appli
   34   553.5    36.6   1613  4  US-10-041-400A-1        Sequence 1, Appli
   35   553.5    36.6   1613  4  US-10-042-091A-1        Sequence 1, Appli
   36     535    35.4   1130  4  US-09-387-375-8        Sequence 8, Appli
   37     535    35.4   1130  4  US-10-041-400A-8        Sequence 8, Appli
   38     535    35.4   1130  4  US-10-042-091A-8        Sequence 8, Appli
   39   530.5    35.1   1378  4  US-09-907-794A-262      Sequence 262, App
   40   530.5    35.1   1378  4  US-09-905-125A-262      Sequence 262, App
   41   530.5    35.1   1378  4  US-09-902-775A-262      Sequence 262, App
   42   530.5    35.1   1378  4  US-09-906-700-262       Sequence 262, App
   43   530.5    35.1   1378  4  US-09-903-603A-262      Sequence 262, App
   44   530.5    35.1   1378  4  US-09-904-920A-262      Sequence 262, App
   45   530.5    35.1   1378  4  US-09-909-064-262       Sequence 262, App


                              ALIGNMENTS

RESULT 1
US-09-016-366A-16
; Sequence 16, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
```

```
;      TELEX:
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1154 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-09-016-366A-16

Alignment Scores:
Pred. No.:              1.56e-156      Length:        1154
Score:                  1512.00        Matches:       275
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     2              Gaps:          0

US-09-598-982C-52 (1-275) x US-09-016-366A-16 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 2
US-08-978-404B-11
; Sequence 11, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1154 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-978-404B-11

Alignment Scores:
Pred. No.:              1.56e-156      Length:        1154
Score:                  1512.00        Matches:       275
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     2              Gaps:          0

US-09-598-982C-52 (1-275) x US-08-978-404B-11 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257
```

```
Qy           81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy          101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy          121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy          141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy          161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy          181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy          201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy          221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy          241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy          261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
                |||||||||||||||||||||||||||||||||||||||||||||
Db          798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 3
US-09-016-366A-18
; Sequence 18, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-18

Alignment Scores:
Pred. No.:              1.93e-144      Length:         1137
Score:                  1402.00        Matches:        252
Percent Similarity:     95.24%         Conservative:   8
Best Local Similarity:  92.31%         Mismatches:     13
Query Match:            92.72%         Indels:         0
DB:                     2              Gaps:           0

US-09-598-982C-52 (1-275) x US-09-016-366A-18 (1-1137)

Qy            3 SerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaPro 22
                :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            3 AATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCA 62

Qy           23 ValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrp 42
                   ||||||||||||:::   |||||||||||||||||||||||||||||||||||||||
Db           63 GGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGG 122

Qy           43 ProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySer 62
                ||||||||||||||||||||||||         |||||||||||||||||||||||||||
Db          123 CCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCC 182

Qy           63 LeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAsp 82
                |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db          183 CTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGAT 242

Qy           83 LeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeu 102
                ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          243 CTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTG 302

Qy          103 ProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIle 122
                ||||||||||||||||||||||||||||||||||||      |||   ||||||||||||
Db          303 CCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATC 362

Qy          123 AlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeu 142
                ||||||||||||||||||||||||||||||:::||||||   ||||||||||||   |||
Db          363 GCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTG 422

Qy          143 ProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAsp 162
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          423 CCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGAT 482

Qy          163 ValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIle 182
                |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db          483 GTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATA 542

Qy          183 MetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspVal 202
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          543 ATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTC 602

Qy          203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGly 222
                ||||||:::|||||||||||||||||||||||||||::::::||||||||||||:::|||
Db          603 CGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGC 662

Qy          223 AspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValVal 242
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          663 GACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTC 722

Qy          243 SerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyr 262
                ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        723 AGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTAC 782

Qy        263 TyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||
Db        783 TACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 821

RESULT 4
US-08-978-404B-13
; Sequence 13, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-13

Alignment Scores:
Pred. No.:              1.93e-144      Length:         1137
Score:                  1402.00        Matches:        252
Percent Similarity:     95.24%         Conservative:   8
Best Local Similarity:  92.31%         Mismatches:     13
Query Match:            92.72%         Indels:         0
DB:                     2              Gaps:           0

US-09-598-982C-52 (1-275) x US-08-978-404B-13 (1-1137)

Qy          3 SerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaPro 22
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCA 62

Qy         23 ValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrp 42
                 ||||||||||||:::   |||||||||||||||||||||||||||||||||||||||
Db         63 GGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGG 122

Qy         43 ProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySer 62
              ||||||||||||||||||||||||         |||||||||||||||||||||||||||
Db        123 CCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCC 182

Qy         63 LeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAsp 82
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        183 CTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGAT 242

Qy         83 LeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeu 102
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTG 302

Qy        103 ProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIle 122
              ||||||||||||||||||||||||||||||||||||      |||   ||||||||||||
Db        303 CCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATC 362

Qy        123 AlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeu 142
              ||||||||||||||||||||||||||||||:::||||||   ||||||||||||   |||
Db        363 GCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTG 422

Qy        143 ProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAsp 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGAT 482

Qy        163 ValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIle 182
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        483 GTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATA 542

Qy        183 MetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspVal 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTC 602

Qy        203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGly 222
              ||||||:::|||||||||||||||||||||||||||::::::||||||||||||:::|||
Db        603 CGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGC 662

Qy        223 AspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValVal 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTC 722

Qy        243 SerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyr 262
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 AGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTAC 782

Qy        263 TyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||
Db        783 TACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 821

RESULT 5
US-09-016-366A-20
; Sequence 20, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-20

Alignment Scores:
Pred. No.:               3.16e-144     Length:        1128
Score:                   1400.00       Matches:       252
Percent Similarity:      94.89%        Conservative:  8
Best Local Similarity:   91.97%        Mismatches:    14
Query Match:             92.59%        Indels:        0
DB:                      2             Gaps:          0

US-09-598-982C-52 (1-275) x US-09-016-366A-20 (1-1128)

Qy          2 LeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAla 21
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCC 61

Qy         22 ProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLys 41
              |||   ||||||||||||:::   ||||||||||||||||||||||||||||||||||||
Db         62 CCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAG 121

Qy         42 TrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGly 61
              |||||||||||||||||||||||||||         ||||||||||||||||||||||||
Db        122 TGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGC 181

Qy         62 SerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLys 81
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        182 TCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTCAAG 241

Qy         82 AspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeu 101
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTG 301

Qy        102 LeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAsp 121
              |||||||||||||||||||||||||||||||||||||||      |||   |||||||||
Db        302 CTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGAC 361

Qy        122 IleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMet 141
              ||||||||||||||||||||||||||||||   :::||||||   ||||||||||||   
Db        362 ATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACC 421

Qy        142 LeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGly 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGC 481

Qy        162 AspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValPro 181
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        482 GATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCC 541

Qy        182 IleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAsp 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGAC 601

Qy        202 ValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLys 221
              |||||||||:::|||||||||||||||||||||||||||::::::||||||||||||:::
Db        602 GTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAG 661

Qy        222 GlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyVal 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTG 721
```

```
Qy        242 ValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThr 261
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACC 781

Qy        262 TyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              ||||||||||||||||||||||||||||||||||||||||||
Db        782 TACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 823

RESULT 6
US-08-978-404B-15
; Sequence 15, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-15

Alignment Scores:
Pred. No.:               3.16e-144     Length:        1128
Score:                   1400.00       Matches:       252
Percent Similarity:      94.89%        Conservative:  8
Best Local Similarity:   91.97%        Mismatches:    14
Query Match:             92.59%        Indels:        0
DB:                      2             Gaps:          0

US-09-598-982C-52 (1-275) x US-08-978-404B-15 (1-1128)

Qy          2 LeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAla 21
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCC 61

Qy         22 ProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLys 41
              |||   ||||||||||||:::   ||||||||||||||||||||||||||||||||||||
Db         62 CCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAG 121

Qy         42 TrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGly 61
              |||||||||||||||||||||||||||         ||||||||||||||||||||||||
```

```
Db         122 TGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGC 181

Qy          62 SerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLys 81
               ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db         182 TCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAG 241

Qy          82 AspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeu 101
               |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 GATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTG 301

Qy         102 LeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAsp 121
               |||||||||||||||||||||||||||||||||||||||      |||   |||||||||
Db         302 CTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGAC 361

Qy         122 IleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMet 141
               ||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db         362 ATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACC 421

Qy         142 LeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGly 161
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGC 481

Qy         162 AspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValPro 181
               ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db         482 GATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCC 541

Qy         182 IleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAsp 201
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 ATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGAC 601

Qy         202 ValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLys 221
               |||||||||:::|||||||||||||||||||||||||||::::::||||||||||||:::
Db         602 GTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAG 661

Qy         222 GlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyVal 241
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 GGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTG 721

Qy         242 ValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThr 261
               |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         722 GTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACC 781

Qy         262 TyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
               ||||||||||||||||||||||||||||||||||||||||||
Db         782 TACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 823

RESULT 7
US-09-016-366A-22
; Sequence 22, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-22

Alignment Scores:
Pred. No.:                1.35e-143      Length:        1081
Score:                    1394.00        Matches:       249
Percent Similarity:       95.88%         Conservative:  7
Best Local Similarity:    93.26%         Mismatches:    11
Query Match:              92.20%         Indels:        0
DB:                       2              Gaps:          0

US-09-598-982C-52 (1-275) x US-09-016-366A-22 (1-1081)

Qy           9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
               ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db           2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy          29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121

Qy          49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy          69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
               |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db         182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy          89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy         109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
               ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db         302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy         129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
               |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db         362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy         149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy         169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy         189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
               ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db         542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy         209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
               ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db         602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661
```

```
Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAAGCCG 802

RESULT 8
US-08-978-404B-17
; Sequence 17, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-17

Alignment Scores:
Pred. No.:                1.35e-143      Length:         1081
Score:                    1394.00        Matches:        249
Percent Similarity:       95.88%         Conservative:   7
Best Local Similarity:    93.26%         Mismatches:     11
Query Match:              92.20%         Indels:         0
DB:                       2              Gaps:           0

US-09-598-982C-52 (1-275) x US-08-978-404B-17 (1-1081)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db          2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121
```

```
Qy         49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy         69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
              |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy         89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy        109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
              ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db        302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy        129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
              |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db        362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy        149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAAGCCG 802

RESULT 9
US-09-917-254-50
; Sequence 50, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Baak, Jan
;  TITLE OF INVENTION: Prognostic Classification of Breast Cancer
;  FILE REFERENCE: B0801/7224(JRV)
;  CURRENT APPLICATION NUMBER: US/09/917,254
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/222,093
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-917-254-50

Alignment Scores:
Pred. No.:                1.35e-143      Length:         1081
Score:                    1394.00        Matches:        249
```

```
Percent Similarity:     95.88%         Conservative:  7
Best Local Similarity:  93.26%         Mismatches:    11
Query Match:            92.20%         Indels:        0
DB:                     4              Gaps:          0

US-09-598-982C-52 (1-275) x US-09-917-254-50 (1-1081)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||:::
Db          2 CTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAGGCCCTGCAGCGA 61

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 121

Qy         49 ArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrp 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGTCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGG 181

Qy         69 ValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVal 88
              |||||||||||||||||||||:::||||||||||||||||||||||||   |||||||||
Db        182 GTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTG 241

Qy         89 GlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIle 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATC 301

Qy        109 ValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGlu 128
              ||||||||||||||||||      |||   ||||||||||||||||||||||||||||||
Db        302 GTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAG 361

Qy        129 GluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThr 148
              |||||||||   :::||||||   ||||||||||||   |||||||||||||||||||||
Db        362 GAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACC 421

Qy        149 PheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPro 168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGC 481

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGT 541

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        542 GACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAC 601

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              ||||||||||||||||||::::::||||||||||||:::|||||||||||||||||||||
Db        602 ATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTG 661

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        662 GTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGT 721

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 781

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
Db        782 CACTATGTCCCCAAAAAGCCG 802

RESULT 10
US-09-079-970A-1
; Sequence 1, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..735
US-09-079-970A-1

Alignment Scores:
Pred. No.:              9.55e-132      Length:        735
Score:                  1284.00        Matches:       227
Percent Similarity:     95.10%         Conservative:  6
Best Local Similarity:  92.65%         Mismatches:    12
Query Match:            84.92%         Indels:        0
DB:                     3              Gaps:          0

US-09-598-982C-52 (1-275) x US-09-079-970A-1 (1-735)

Qy         31 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 50
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         51 ArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 70
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy         71 ThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgValGlnLeu 90
              |||||||||||||||:::||||||||||||||||||||||||   |||||||||||||||
Db        121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         91 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy        111 ProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 130
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        131 ValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThrPhePro 150
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360
```

```
Qy          151 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluProLeuPro 170
                |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db          361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy          171 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 190
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy          191 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAspMetLeu 210
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy          211 CysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValCys 230
                ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db          541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy          231 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCysAlaGln 250
                ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db          601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy          251 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy          271 ValProLysLysPro 275
                |||||||||||||||
Db          721 GTCCCCAAAAAGCCG 735

RESULT 11
US-09-079-970A-4
; Sequence 4, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
```

```
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  7..753
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7..18
US-09-079-970A-4

Alignment Scores:
Pred. No.:              1.03e-131      Length:          771
Score:                  1284.00        Matches:         227
Percent Similarity:     95.10%         Conservative:    6
Best Local Similarity:  92.65%         Mismatches:      12
Query Match:            84.92%         Indels:          0
DB:                     3              Gaps:            0

US-09-598-982C-52 (1-275) x US-09-079-970A-4 (1-771)

Qy           31 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 50
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78

Qy           51 ArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 70
                         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138

Qy           71 ThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgValGlnLeu 90
                |||||||||||||||:::||||||||||||||||||||||||   |||||||||||||||
Db          139 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198

Qy           91 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 110
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258

Qy          111 ProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 130
                ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db          259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318

Qy          131 ValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThrPhePro 150
                |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db          319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378

Qy          151 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluProLeuPro 170
                |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db          379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438

Qy          171 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 190
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498

Qy          191 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAspMetLeu 210
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558

Qy          211 CysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValCys 230
                ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db          559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618

Qy          231 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCysAlaGln 250
                ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db          619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678

Qy          251 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738

Qy          271 ValProLysLysPro 275
                |||||||||||||||
Db          739 GTCCCCAAAAAGCCG 753

RESULT 12
US-08-978-404B-1
```

```
; Sequence 1, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-1

Alignment Scores:
Pred. No.:                5.09e-115      Length:         1031
Score:                    1134.00        Matches:        203
Percent Similarity:       82.78%         Conservative:   23
Best Local Similarity:    74.36%         Mismatches:     45
Query Match:              75.00%         Indels:         2
DB:                       2              Gaps:           1

US-09-598-982C-52 (1-275) x US-08-978-404B-1 (1-1031)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||   ||||||||||||   ||||||:::|||:::|||      :::|||||||||
Db         27 ATGCTAAAGCTGCTGCTGCTCACGCTGCCCCTCCTGTCCAGCCTGGTGCATGCAGCCCCC 86

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                 |||      |||:::   :::   ||||||||||||||||||||||||      :::
Db         87 GGTCCA------GCTATGACACGAGAAGGCATTGTGGGGGACAGGAGGCACATGGGAAC 140

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              |||||||||||||||||||||||||||      |||   |||||||||||||||||||||
Db        141 AAGTGGCCCTGGCAGGTGAGCCTGCGTGCCAATGACACCTACTGGATGCATTTCTGCGGT 200

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        201 GGCTCCCTCATCCACCCACAGTGGGTGCTCACTGCGGCACACTGTGTGGGACCGGATGTT 260

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                 |||         :::|||||||||||||||:::|||:::|||||||||   |||   
Db        261 GCTGACCCCAACAAGGTCAGAGTACAGCTCCGTAAGCAGTACCTCTATTACCATGACCAC 320

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              |||:::   ||||||:::||||||   ||||||   |||||||||:::|||   ||||||
Db        321 CTGATGACTGTGAGCCAGATCATCACACACCCCGACTTCTACATCGTCCAGGATGGGGCA 380

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||:::|||      |||||||||||||||      ||||||   |||
Db        381 GACATTGCCCTGCTGAAACTCACAAACCCTGTGAACATTTCTGACTATGTCCACCCTGTC 440

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||   |||      ||||||||||||||||||
Db        441 CCCCTACCTCCTGCCTCAGAGACCTTCCCCTCAGGAACGTTGTGCTGGGTGACAGGCTGG 500

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||::::::||||||         ||||||||||||||||||||||||:::|||:::|||
Db        501 GGTAACATCGACAATGGTGTAAACCTGCCGCCACCATTTCCTTTGAAGGAGGTGCAAGTT 560

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||:::|||||||||:::||||||   |||||||||   |||      |||||||||
Db        561 CCCATTATAGAAAACCACCTTTGTGACTTGAAGTATCACAAAGGTCTCATCACAGGTGAC 620

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              :::|||   |||:::|||||||||||||||||||||||||||         |||||||||
Db        621 AATGTCCACATTGTCCGAGATGACATGCTGTGTGCTGGGAATGAAGGACATGACTCCTGC 680

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||      ||||||||||||||||||
Db        681 CAGGGCGACTCCGGAGGACCTCTGGTCTGCAAGGTAGAAGACACCTGGCTGCAGGCAGGC 740

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        741 GTGGTCAGCTGGGGTGAGGGCTGTGCACAGCCCAACAGGCCTGGCATCTACACCCGGGTC 800

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLys 273
              |||||||||||||||||||||||||||||||||||||||
Db        801 ACCTATTACTTGGACTGGATCCACCACTATGTCCCCAAG 839

RESULT 13
US-09-016-366A-14
; Sequence 14, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-14

Alignment Scores:
Pred. No.:              7.32e-115     Length:       1108
Score:                  1133.00       Matches:      198
Percent Similarity:     83.70%        Conservative: 28
Best Local Similarity:  73.33%        Mismatches:   42
Query Match:            74.93%        Indels:       2
DB:                     2             Gaps:         1

US-09-598-982C-52 (1-275) x US-09-016-366A-14 (1-1108)

Qy          4 LeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProVal 23
              |||||||||   ||||||   :::|||||||||      |||:::||||||   |||
Db         52 CTGCTGCTGTGGGCACTGTCCCTCCTGGCTAGTCTGGTGTACTCAGCCCCTCGCCCA--- 108

Qy         24 GlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpPro 43
                 |||   |||:::   |||||||||||||||   ||||||      ||||||||||||
Db        109 ---GCCAATCAGCGAGTGGGCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCC 165

Qy         44 TrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeu 63
              ||||||||||||||||||   :::      ||||||:::|||||||||||||||||||||
Db        166 TGGCAGGTGAGCCTGAGATTTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTC 225

Qy         64 IleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeu 83
              ||||||||||||||||||||||||||||||||||||:::||||||   :::|||
Db        226 ATCCACCCACAGTGGGTGCTCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCA 285

Qy         84 AlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuPro 103
                       |||||||||||||||||||||:::|||||||||   ||||||||||||
Db        286 CAGCTCTTCCGGGTGCAGCTTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCT 345

Qy        104 ValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAla 123
              ::::::||||||:::|||||||||   :::|||      :::   |||||||||:::|||
Db        346 TTGAACCGGATCGTGGTGCACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCC 405

Qy        124 LeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeuPro 143
              |||||||||||||||   |||||||||:::|||:::   :::|||   :::   ||||||
Db        406 CTGCTGGAGCTTGAGGTCCCTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCC 465

Qy        144 ProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspVal 163
              |||||||||||||||||||||||||||      ||||||||||||||||||||||||:::
Db        466 CCTGCCTCGGAGACCTTCCCCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATT 525

Qy        164 AspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIleMet 183
              |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||:::
Db        526 GATAATGACGAGCCTCTCCCACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTG 585

Qy        184 GluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArg 203
              ||||||   :::||||||   |||||||||   |||   |||||||||||||||
Db        586 GAAAACAGCCTGTGTGACCGGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCC 645

Qy        204 IleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAsp 223
              |||:::   |||   ||||||||||||||||||::::::||||||||||||:::||||||
Db        646 ATTGTCCATGATGGCATGCTGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGAT 705

Qy        224 SerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSer 243
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        706 TCAGGGGGGCCACTGGTCTGCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGC 765

Qy        244 TrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyr 263
              |||   |||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        766 TGGGGTGAGGGCTGCGCACAGCCCAACAAGCCTGGCATCTACACCCGGGTGACATACTAC 825

Qy        264 LeuAspTrpIleHisHisTyrValProLys 273
              |||||||||||||||   |||||||||:::
Db        826 TTAGACTGGATCCACCGCTATGTCCCTGAG 855
```

```
RESULT 14
US-08-978-404B-20
; Sequence 20, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-20

Alignment Scores:
Pred. No.:              7.32e-115     Length:       1108
Score:                  1133.00       Matches:      198
Percent Similarity:     83.70%        Conservative: 28
Best Local Similarity:  73.33%        Mismatches:   42
Query Match:            74.93%        Indels:       2
DB:                     2             Gaps:         1

US-09-598-982C-52 (1-275) x US-08-978-404B-20 (1-1108)

Qy          4 LeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProVal 23
              |||||||||   ||||||   :::|||||||||      |||:::||||||   |||
Db         52 CTGCTGCTGTGGGCACTGTCCCTCCTGGCTAGTCTGGTGTACTCAGCCCCTCGCCCA--- 108

Qy         24 GlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpPro 43
                 |||   |||:::   |||||||||||||||   ||||||      ||||||||||||
Db        109 ---GCCAATCAGCGAGTGGGCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCC 165

Qy         44 TrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeu 63
              ||||||||||||||||||   :::      ||||||:::|||||||||||||||||||||
Db        166 TGGCAGGTGAGCCTGAGATTTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTC 225

Qy         64 IleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeu 83
              ||||||||||||||||||||||||||||||||||||:::||||||   :::|||
Db        226 ATCCACCCACAGTGGGTGCTCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCA 285

Qy         84 AlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuPro 103
                       |||||||||||||||||||||:::|||||||||   ||||||||||||
```

```
Db        286 CAGCTCTTCCGGGTGCAGCTTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCT 345

Qy        104 ValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAla 123
              ::::::||||||:::|||||||||   :::|||      :::   |||||||||:::|||
Db        346 TTGAACCGGATCGTGGTGCACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCC 405

Qy        124 LeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeuPro 143
              |||||||||||||||   |||||||||:::|||:::   :::|||   :::   ||||||
Db        406 CTGCTGGAGCTTGAGGTCCCTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCC 465

Qy        144 ProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspVal 163
              |||||||||||||||||||||||||||      ||||||||||||||||||||||||:::
Db        466 CCTGCCTCGGAGACCTTCCCCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATT 525

Qy        164 AspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIleMet 183
              |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||:::
Db        526 GATAATGACGAGCCTCTCCCACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTG 585

Qy        184 GluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArg 203
              ||||||   :::||||||   |||||||||   |||   |||||||||||||||
Db        586 GAAAACAGCCTGTGTGACCGGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCC 645

Qy        204 IleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAsp 223
              |||:::   |||   ||||||||||||||||||::::::||||||||||||:::||||||
Db        646 ATTGTCCATGATGGCATGCTGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGAT 705

Qy        224 SerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSer 243
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        706 TCAGGGGGGCCACTGGTCTGCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGC 765

Qy        244 TrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyr 263
              |||   |||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        766 TGGGGTGAGGGCTGCGCACAGCCCAACAAGCCTGGCATCTACACCCGGGTGACATACTAC 825

Qy        264 LeuAspTrpIleHisHisTyrValProLys 273
              |||||||||||||||   |||||||||:::
Db        826 TTAGACTGGATCCACCGCTATGTCCCTGAG 855

RESULT 15
US-08-845-998-3
; Sequence 3, Application US/08845998
; Patent No. 5879892
;  GENERAL INFORMATION:
;    APPLICANT:  Van Baren, Nicolas
;    APPLICANT:  Coulie, Pierre G.
;    APPLICANT:  De Smet, Charles
;    APPLICANT:  Lucas, Sophie
;    APPLICANT:  Boon, Thierry
;    TITLE OF INVENTION:  LEUKEMIA ASSOCIATED GENES
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,998
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Van Amsterdam, John R.
;      REGISTRATION NUMBER:  40,212
;      REFERENCE/DOCKET NUMBER:  L0461/7008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)720-3500
;      TELEFAX:  (617)720-2441
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2259 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  8..577
US-08-845-998-3

Alignment Scores:
Pred. No.:              8.51e-113       Length:         2259
Score:                  1118.50         Matches:        236
Percent Similarity:     41.68%          Conservative:   17
Best Local Similarity:  38.88%          Mismatches:     22
Query Match:            73.97%          Indels:         333
DB:                     2               Gaps:           2

US-09-598-982C-52 (1-275) x US-08-845-998-3 (1-2259)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||   ||||||   ||||||
Db          8 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGTGGCCCCT 67

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   |||||||||||||||   |||||||||||||||||||||||||||||||||
Db         68 GCCCCAGGCCAGGCCCTGCAGCAAACGGGCATTGTTGGGGGGCAGGAGGCCCCCAGGAGC 127

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              |||||||||||||||||||||||||||||||||      |||||||||||||||||||||
Db        128 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGGCCCATACTGGATGCACTTCTGCGGG 187

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::   ||||||:::
Db        188 GGCTCCCTCATCCACCCCCAGTGGGTGCTAACCGCGGCGCACTGCGTGGAACCGGACATC 247

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        248 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 307

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACCGGGGCG 367

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||||||||||||||   :::|||||||||
Db        368 GACATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCACATCCACACGGTC 427

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 ACGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 487

Qy        161 GlyAspValAspAsnAsp------------------------------------------ 166
              |||||||||||||||
Db        488 GGCGACGTGGACAATAA-TGGTGGGTGTTGGGGACAGCGGGAGGCCGGGCCAGGTGGGCA 546

Qy        166 ------------------------------------------------------------ 166

Db        547 CCAAGTCACAGCCACAGGCCAGTCCGTGGGGTGACAGGGTCCCTCAGGGCGGCTCAGGGA 606

Qy        166 ------------------------------------------------------------ 166

Db        607 GGGGGACTGTGGAGGCCAGGATGGATGGAGCAGGCGGTGGCGAGAGGCAGCAGGTGCCCT 666

Qy        166 ------------------------------------------------------------ 166

Db        667 GAGCAGAGACGGTGAGTCCAAAGGGCCTGGGCGTCCCCCACCCCAGGGGTTTGGAGAGTC 726

Qy        166 ------------------------------------------------------------ 166
```

```
Db        727 CCTTAGCACCTCCGTGCCTCGGTTTCCCCTTGCCTGAAAGGGTGCATCAAAAGTTTGTAC 786

Qy        166 ------------------------------------------------------------ 166

Db        787 GTCACGGACTTGCTATGTGGAGAGAGAAATCACACGGGGGTCTTGCTGGAAGGAGAGAGA 846

Qy        166 ------------------------------------------------------------ 166

Db        847 CCGGTGCTGGGATGAGACCTGCCTGCCCTCCATCCCTGTGCTACAGACAAGGCAGGGGCC 906

Qy        166 ------------------------------------------------------------ 166

Db        907 TGGGAATCGGGGTCGTGGCAGTGCTGTGGGGGGCTGGACGAAGCTCACTGTGGCCCTCCA 966

Qy        166 ------------------------------------------------------------ 166

Db        967 CGAGGCACATTTTCACTTCTAGAAGGTCTTGTCCCCATTTTATCCACAATTCAGAGCAAA 1026

Qy        166 ------------------------------------------------------------ 166

Db       1027 GCTTTGGGGTACAGCCTGACGCAACCCTGGGCTGTGACCTCTGGGTCACTCCAGAAGGGG 1086

Qy        166 ------------------------------------------------------------ 166

Db       1087 CCTGAGCCACTGTCCCGCTATTCCGCCCCACACAGCGGGGAAGCTGAGCCCAGCGCCCTG 1146

Qy        166 ------------------------------------------------------------ 166

Db       1147 TGTTCCCCTCGGCTAGGGCCAACCGTGGACCATGGGCCTAGCCCAGACGAAAGTCAGCTG 1206

Qy        167 ---------------------------------------------GluPro--------- 168
                                                           ||||||
Db       1207 AGCCCAGGGGGAGACACGGGTCGGGCTCTGCACCCCCGTGCCATGGAGCCCAGCTTGGCA 1266

Qy        168 ------------------------------------------------------------ 168

Db       1267 ACCTCCAGGGCCCTCCCCTCCCTTCCCCAGATGGGGCTTAAATGAGGCCAGGGACCCAGG 1326

Qy        168 ------------------------------------------------------------ 168

Db       1327 ACCAGCCTCAGCGGAGGGGCCTGGACTGCATTCACCGCCCCTTCCCCGGGGCTGCAGGCA 1386

Qy        168 ------------------------------------------------------------ 168

Db       1387 CAGAACAGCACTGGGCCCATGGTGCCATCTCCCCTGCCCGTGACTCTGCCACCAAGTCCA 1446

Qy        168 ------------------------------------------------------------ 168

Db       1447 CGAAGCAGCACCCAGCCGGCCCCAGACCCGGCTCCACGCCCCCCTCCGCCCCCAGTGCAC 1506

Qy        169 LeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCys 188
              ||||||||||||:::|||||||||:::|||:::||||||::::::|||||||||:::|||
Db       1507 CTGCCGCCGCCATACCCGCTGAAGGAGGTGGAAGTCCCCGTAGTGGAAAACCACCTTTGC 1566

Qy        189 AspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAsp 208
              :::|||:::||||||   |||   :::||||||         :::|||:::|||||||||
Db       1567 AACGCGGAATATCACACCGGCCTCCATACGGGCCACAGCTTTCAAATCGTCCGCGATGAC 1626

Qy        209 MetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeu 228
              |||||||||||||||:::         |||||||||:::|||||||||||||||||||||
Db       1627 ATGCTGTGTGCGGGGAGCGAAAATCACGACTCCTGCCAGGGTGACTCTGGAGGGCCCCTG 1686

Qy        229 ValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGluGlyCys 248
              |||||||||||||||||||||   |||||||||   ||||||||||||:::|||   |||
Db       1687 GTCTGCAAGGTGAATGGCACCTAACTGCAGGCGGCCGTGGTCAGCTGGGAGGAGAGCTGT 1746

Qy        249 AlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHis 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 GCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC 1806

Qy        269 HisTyrValProLysLysPro 275
              |||||||||||||||||||||
```

```
Db       1807 CACTATGTCCCCAAGAAGCCC 1827

Search completed: August 27, 2005, 23:44:31
Job time : 194 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         August 30, 2005, 16:09:21 ; Search time 6409 Seconds
                                           (without alignments)
                                           254.007 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:      8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_29082005_080947_26125/app_query.fasta_1.455
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_708_@runat_29082005_080947_26125 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_16Dec04:*
                1:  geneseqn1980s:*
                2:  geneseqn1990s:*
                3:  geneseqn2000s:*
                4:  geneseqn2001as:*
                5:  geneseqn2001bs:*
                6:  geneseqn2002as:*
                7:  geneseqn2002bs:*
                8:  geneseqn2003as:*
                9:  geneseqn2003bs:*
                10:  geneseqn2003cs:*
                11:  geneseqn2003ds:*
                12:  geneseqn2004as:*
                13:  geneseqn2004bs:*

          Pred. No. is the number of results predicted by chance to have a
          score greater than or equal to the score of the result being printed,
          and is derived by analysis of the total score distribution.

                                     SUMMARIES
                     %
Result               Query
  No.     Score  Match Length DB  ID                      Description
 ------------------------------------------------------------------------
     1     1512  100.0   1154  2  AAV44328                Aav44328 Human mas
     2     1512  100.0   1154  2  AAV42710                Aav42710 Human mas
     3     1512  100.0   1154  3  AAA34959                Aaa34959 Human ade
     4     1512  100.0   1154  3  AAF21081                Aaf21081 Human low
     5     1512  100.0   1154 10  ABZ96775                 Abz96775 Human nuc
```

```
     6     1512  100.0   1154 10  ADK52587                Adk52587 Hematolog
     7     1512  100.0   1154 11  ABD20624                Abd20624 Human pul
     8     1512  100.0   1158  9  ACH04035               Ach04035 Human cDN
     9     1512  100.0   1158 12  ADP10384                Adp10384 Reference
    10     1512  100.0   1158 13  ADP56069                Adp56069 Human PRO
    11     1512  100.0  17133  3  AAA34962               Aaa34962 Human ade
    12     1512  100.0  17133  3  AAF21084               Aaf21084 Human low
    13     1512  100.0  17133 10  ABZ96778                Abz96778 Human nuc
    14     1512  100.0  17133 11  ABD20627                Abd20627 Human pul
    15     1430   94.6   1142 13  ADQ85402                Adq85402 Human tum
    16     1416   93.7   1334 12  ADQ23979                Adq23979 Human sof
    17     1405   92.9   1143  3  AAA34955               Aaa34955 Human ade
    18     1405   92.9   1143  3  AAF21077               Aaf21077 Human low
    19     1405   92.9   1143 10  ABZ96771                Abz96771 Human nuc
    20     1405   92.9   1143 10  ADK52589                Adk52589 Hematolog
    21     1405   92.9   1143 11  ABD20620                Abd20620 Human pul
    22     1405   92.9   1143 12  ADJ60631                Adj60631 Tryptase-
    23     1405   92.9   1143 12  ADO46120                Ado46120 Human oli
    24     1405   92.9   1145  3  AAA34956               Aaa34956 Human ade
    25     1405   92.9   1145  3  AAF21078               Aaf21078 Human low
    26     1405   92.9   1145 10  ABZ96772                Abz96772 Human nuc
    27     1405   92.9   1145 10  ADJ94865                Adj94865 Novel NOV
    28     1405   92.9   1145 11  ABD20621                Abd20621 Human pul
    29     1405   92.9  42587 12  ADJ61642                Adj61642 Concateme
    30     1402   92.7   1137  2  AAV44329               Aav44329 Human mas
    31     1402   92.7   1137  2  AAV42711               Aav42711 Human mas
    32     1402   92.7   1137  3  AAA34957               Aaa34957 Human ade
    33     1402   92.7   1137  3  AAF21079               Aaf21079 Human low
    34     1402   92.7   1137 10  ABZ96773                Abz96773 Human nuc
    35     1402   92.7   1137 11  ABD20622                Abd20622 Human pul
    36     1401   92.7   2662  6  AAD23854               Aad23854 Human pro
    37     1400   92.6   1128  2  AAV44330               Aav44330 Human mas
    38     1400   92.6   1128  2  AAV42712               Aav42712 Human mas
    39     1394   92.2   1081  2  AAV44331               Aav44331 Human mas
    40     1394   92.2   1081  2  AAV42713               Aav42713 Human mas
    41     1394   92.2   1081  3  AAA34960               Aaa34960 Human ade
    42     1394   92.2   1081  3  AAF21082               Aaf21082 Human low
    43     1394   92.2   1081  6  ABL62377               Abl62377 Colon ade
    44     1394   92.2   1081  6  ABL66816               Abl66816 Lung canc
    45     1394   92.2   1081  6  ABL61818               Abl61818 Colon ade


                                  ALIGNMENTS

RESULT 1
AAV44328
ID   AAV44328 standard; DNA; 1154 BP.
XX
AC   AAV44328;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human mast cell tryptase alpha nucleic acid sequence.
XX
KW   Mast cell tryptase alpha; human; MCP-7; mast cell protease 7; blood clot;
KW   anticoagulant; myocardial infarction; reocclusion; thromboembolism;
KW   cerebral embolism; thrombosis; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             18. .845
FT                   /*tag=  a
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
```

```
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
DR   P-PSDB; AAW64237.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Disclosure; Page 63; 92pp; English.
XX
CC   This nucleotide sequence includes a coding region for human mast cell
CC   tryptase alpha (see AAW64237). The invention provides: compositions
CC   comprising an isolated tryptase-7 that may include chimeric proteins that
CC   contain (a) a human tryptase for all but the active site region and (b)
CC   the substrate-binding pocket of mouse tryptase-7 or its homologues (see
CC   AAW64233-36); a method for treating a blood clot by administering a
CC   nucleic acid molecule that codes for a tryptase-7, or an expression
CC   product, to decrease fibrinogen activity; a nucleic acid encoding a
CC   serine protease (SP); and a method of producing a mature SP by expressing
CC   the inactive zymogen in a host cell, and cleaving the enterokinase
CC   susceptibility domain. The tryptase-7 polypeptides can be used to treat
CC   disorders mediated by undesirable thrombus clot formation such as
CC   myocardial infarction and reocclusion following angioplasty of blood
CC   clots associated with pulmonary thromboembolism, deep vein thrombosis,
CC   cerebral embolism, renal vein and peripheral arterial thrombosis. They
CC   are also useful for all surgical procedures that require decreased blood
CC   clots
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.18e-129     Length:        1154
Score:                  1512.00       Matches:       275
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     2             Gaps:          0

US-09-598-982C-52 (1-275) x AAV44328 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 2
AAV42710
ID   AAV42710 standard; cDNA; 1154 BP.
XX
AC   AAV42710;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human mast cell tryptase alpha encoding cDNA.
XX
KW   Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
KW   tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
KW   antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
KW   hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
KW   inflammatory skin condition; human; mast cell tryptase alpha; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             18..845
FT                   /*tag=  a
FT                   /product= "Human mast cell tryptase alpha"
XX
PN   WO9833812-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US001865.
XX
PR   05-FEB-1997;   97US-0037090P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL,  Huang C;
XX
DR   WPI; 1998-437390/37.
DR   P-PSDB; AAW63173.
XX
PT   Tryptase-6 complex inhibitory peptides - used to treat mast cell-mediated
PT   inflammatory disorders e.g. asthma.
XX
PS   Disclosure; Page 44-45; 69pp; English.
XX
CC   This cDNA encodes the human mast cell tryptase alpha which is a homologue
CC   of the mouse mast cell protease (mMCP-6) zymogen. The invention provides
CC   sequences shown in AAW63160 to AAW63169 that are inhibitors of mMCP-6.
CC   These peptides which are tryptase-6 complex inhibitors, can be used for
CC   treating a mast cell-mediated inflammatory disorder. The inhibitors can
```

```
CC   be used to treat inflammatory disorders including asthma, allergic
CC   rhinitis, urticaria and antioedema, eczematous dermatitis (atopic
CC   dermatitis), anaphylaxis, hyperproliferative skin disease, peptic ulcers,
CC   inflammatory bowel disorder, hyperresponsiveness and inflammatory skin
CC   conditions
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.18e-129      Length:       1154
Score:                  1512.00        Matches:      275
Percent Similarity:     100.00%        Conservative: 0
Best Local Similarity:  100.00%        Mismatches:   0
Query Match:            100.00%        Indels:       0
DB:                     2              Gaps:         0

US-09-598-982C-52 (1-275) x AAV42710 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842
```

```
RESULT 3
AAA34959
ID   AAA34959 standard; DNA; 1154 BP.
XX
AC   AAA34959;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human adenosine receptor related polynucleotide SEQ ID NO:2648.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US017712.
XX
PR   03-AUG-1998;   98US-0095212P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstruction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers.
XX
PS   Disclosure; Page 808-809; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC   sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC   from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
CC   AAA33992) are specifically claimed ONs from the present invention. N.B.
CC   Sequences given in the disclosure of the present invention do not match
CC   up with their corresponding SEQ ID NO: sequences given in the sequence
CC   listing
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.18e-129      Length:       1154
```

```
Score:                  1512.00        Matches:       275
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     3              Gaps:          0

US-09-598-982C-52 (1-275) x AAA34959 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 4
AAF21081
ID   AAF21081 standard; DNA; 1154 BP.
XX
AC   AAF21081;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2648.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US008020.
XX
PR   06-APR-1999;   99US-0127958P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not trigger
PT   adenosine receptors during metabolism, useful e.g. for treating cancers
PT   and respiratory obstructions.
XX
PS   Disclosure; Page 881; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC   surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.18e-129      Length:        1154
Score:                  1512.00        Matches:       275
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
```

```
DB:                       3               Gaps:          0

US-09-598-982C-52 (1-275) x AAF21081 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 5
ABZ96775
ID   ABZ96775 standard; DNA; 1154 BP.
XX
AC   ABZ96775;
XX
DT   17-OCT-2003  (first entry)
XX
DE   Human nucleic acid sequence.
XX
KW   Human; antisense; lung dysfunction; nasal airway dysfunction;
KW   antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW   antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
```

```
KW   antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW   adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW   lung inflammation; respiratory disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285308-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013135.
XX
PR   24-APR-2001; 2001US-0286137P.
XX
PA   (EPIG-) EPIGENESIS PHARM INC.
XX
PI   Nyce JW,  Li Y,  Sandrasagra A,  Katz E,  Pabalan J,  Aguilar D;
PI   Miller S,  Tang L,  Shahabuddin S;
XX
DR   WPI; 2003-229219/22.
XX
PT   Pharmaceutical composition for treating ailments associated with impaired
PT   respiration, has oligo(s) antisense to specific gene(s) or its
PT   corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT   ubiquinone.
XX
PS   Disclosure; SEQ ID NO 12017; 872pp; English.
XX
CC   The invention relates to a novel pharmaceutical composition, which has a
CC   first active agent comprising an oligonucleotide antisense to the
CC   initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC   5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC   junctions of genes encoding a polypeptide associated with lung and/or
CC   nasal airway dysfunction and a second active agent comprising an
CC   antiinflammatory steroid and ubiquinone. A composition of the invention
CC   has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
CC   immunosuppressive, and cytostatic activity. The composition may have a
CC   use in antisense gene therapy. The composition is useful for treating or
CC   preventing a respiratory, lung or malignant disease or condition, also
CC   for enhancing the prophylactic or therapeutic respiratory effect of an
CC   antiinflammatory steroid in a subject, for reducing or depleting levels
CC   of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC   receptor, producing bronchodilation, increasing levels of ubiquinone or
CC   lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC   lung inflammation, lung allergies, or a respiratory disease or condition.
CC   Note: The sequence data for this patent is not represented in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.18e-129      Length:        1154
Score:                  1512.00        Matches:       275
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     10             Gaps:          0

US-09-598-982C-52 (1-275) x ABZ96775 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197
```

```
Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 6
ADK52587
ID   ADK52587 standard; DNA; 1154 BP.
XX
AC   ADK52587;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Hematological disorder associated Gene ID 1847.
XX
KW   cytostatic; antianemic; antisickling; virucide; hemostatic; nephrotropic;
KW   cytostatic; thrombolytic; antiparasitic; gene therapy;
KW   hematologic disorder; cancer; Sickle Cell Anemia;
KW   Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
KW   Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
KW   transfusion reaction; Erythroblastosis; mechanical trauma;
KW   micro-angiopathic hemolytic anemia; parasite infection; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             18. .845
FT                   /*tag=  a
XX
PN   WO2003065871-A2.
XX
PD   14-AUG-2003.
XX
PF   28-JAN-2003; 2003WO-US002484.
XX
PR   04-FEB-2002; 2002US-0354333P.
PR   28-FEB-2002; 2002US-0360258P.
PR   15-MAR-2002; 2002US-0364476P.
PR   26-APR-2002; 2002US-0375626P.
PR   06-JUN-2002; 2002US-0386494P.
PR   24-JUN-2002; 2002US-0390965P.
PR   28-JUN-2002; 2002US-0392480P.
PR   03-JUL-2002; 2002US-0394128P.
PR   31-JUL-2002; 2002US-0399783P.
PR   13-AUG-2002; 2002US-0403221P.
PR   30-AUG-2002; 2002US-0407045P.
PR   25-NOV-2002; 2002US-0429048P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Carroll JM,  Healy A,  Weich NS,  Kelly LM;
XX
DR   WPI; 2003-731464/69.
DR   P-PSDB; ADK52588.
XX
PT   Identifying a compound capable of treating a hematologic disorder (e.g.
PT   anemia or leukemia) comprises assaying the ability of the compound to
PT   modulate the expression or activity of e.g. 131,148, 199 or 12303
PT   polypeptide or nucleic acid.
XX
PS   Disclosure; SEQ ID NO 45; 232pp; English.
XX
CC   The invention relates to a method of identifying a compound capable of
CC   treating a hematologic disorder comprises assaying the ability of the
CC   compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
CC   194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
CC   13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
CC   acid expression or polypeptide activity, thus, identifying a compound
CC   capable of treating a hematologic disorder. The methods are useful in
CC   diagnosing, preventing and treating hematological disorders, such as
CC   cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC   Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders
CC   associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC   antibody-mediated disorders such as transfusion reactions and
CC   Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC   angiopathic hemolytic anemias, infections by parasites or chemical
CC   injuries. The methods may also be used for identifying compounds that
CC   modulate hematological disorders. This sequence corresponds to one of the
CC   genes modulated the compound.
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.18e-129       Length:        1154
Score:                  1512.00         Matches:       275
Percent Similarity:     100.00%         Conservative:  0
Best Local Similarity:  100.00%         Mismatches:    0
Query Match:            100.00%         Indels:        0
DB:                     10              Gaps:          0

US-09-598-982C-52 (1-275) x ADK52587 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257
```

```
Qy          81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy         101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy         121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy         141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy         161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy         181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy         201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy         221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy         241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy         261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
               |||||||||||||||||||||||||||||||||||||||||||||
Db         798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 7
ABD20624
ID   ABD20624 standard; DNA; 1154 BP.
XX
AC   ABD20624;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human pulmonary and inflammatory target DNA #235.
XX
KW   Human; antisense; bronchoconstriction; allergy; hyposecretion; pain;
KW   respiratory tract inflammation; adenosine sensitivity; lung; cancer;
KW   surfactant depletion; antiallergic; antiinflammatory; antiasthmatic;
KW   analgesic; hypotensive; immunosuppressive; cytostatic; cystic fibrosis;
KW   beta-adrenergic agonist; respiratory disease; pulmonary vasoconstriction;
KW   respiratory distress syndrome; allergic rhinitis; pulmonary hypertension;
KW   emphysema; chronic obstructive pulmonary disease; cancer; bronchitis;
KW   pulmonary transplantation rejection; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285309-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013143.
XX
PR   24-APR-2001; 2001US-0286036P.
XX
PA   (EPIG-) EPIGENESIS PHARM INC.
XX
PI   Nyce JW,  Li Y,  Sandrasagra A,  Katz E,  Pabalan J,  Aguilar D;
PI   Miller S,  Tang L,  Shahabuddin S;
XX
DR   WPI; 2003-093058/08.
XX
PT   Pharmaceutical composition for treating asthma, has antisense
PT   oligonucleotide containing less percentage of adenosine, targeted to
PT   nucleic acids associated with lung airway or lung dysfunction, and
PT   bronchodilating agent.
XX
PS   Claim 15; SEQ ID NO 12017; 763pp; English.
XX
CC   This invention describes a novel composition (a) a first active agent,
CC   comprising oligonucleotides, effective for alleviating
CC   bronchoconstriction, respiratory tract inflammation, allergies and
CC   reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors,
CC   surfactant depletion or hyposecretion, when administered to a mammal. The
CC   oligonucleotides are derived from a gene encoding or regulating
CC   expression of a target polypeptide associated with lung airway or lung
CC   dysfunction or cancer and can be anti-sense to the corresponding mRNA.
CC   The invention also describes a kit, that comprises: (a) a delivery
CC   device, in separate containers, (b) the oligonucleotides, (c)
CC   instructions for adding a carrier and for use of the kit. The composition
CC   of the invention has antiallergic, antiinflammatory, antiasthmatic,
CC   analgesic, hypotensive, immunosuppressive and cytostatic activity, is a
CC   beta-adrenergic agonist. The composition is useful for preventing or
CC   treating a respiratory, lung or malignant disease. The administered
CC   composition comprises oligo and is administered to reduce the production
CC   or availability, or to increase the degradation of the target mRNA or to
CC   reduce the amount of target polypeptide present in the lungs. The
CC   pulmonary obstruction, and/or bronchoconstriction and/or lung
CC   inflammation, allergies and/or surfactant hypoproduction are associated
CC   with a disease or condition such as pulmonary vasoconstriction,
CC   inflammation, allergies, asthma, impeded respiration, respiratory
CC   distress syndrome, pain, cystic fibrosis, allergic rhinitis, pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary
CC   transplantation rejection, pulmonary infections, bronchitis or cancer.
CC   The reduced adenosine content of the anti-sense oligos corresponding to
CC   thymidines present in the target RNA serves to prevent the breakdown of
CC   the oligonucleotides into products that free adenosine into the system
CC   e.g., lung, brain, heart, kidney, etc, tissue environment and thereby, to
CC   prevent any unwanted effects due to it
XX
SQ   Sequence 1154 BP; 203 A; 411 C; 330 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:             5.18e-129     Length:        1154
Score:                 1512.00       Matches:       275
Percent Similarity:    100.00%       Conservative:  0
Best Local Similarity: 100.00%       Mismatches:    0
Query Match:           100.00%       Indels:        0
DB:                    11            Gaps:          0

US-09-598-982C-52 (1-275) x ABD20624 (1-1154)

Qy           1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy          21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy          41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy          61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy          81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317
```

```
Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 8
ACH04035
ID   ACH04035 standard; cDNA; 1158 BP.
XX
AC   ACH04035;
XX
DT   26-SEP-2003  (first entry)
XX
DE   Human cDNA differentially expressed in lung cancer #240.
XX
KW   Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
KW   respiratory disorder; lung cancer; asthma; human.
XX
OS   Homo sapiens.
XX
PN   US2003065157-A1.
XX
PD   03-APR-2003.
XX
PF   04-APR-2002; 2002US-00116802.
XX
PR   04-APR-2001; 2001US-0281593P.
XX
PA   (LASE/) LASEK A W.
XX
PI   Lasek AW;
XX
DR   WPI; 2003-540803/51.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   respiratory disorders, useful for diagnosing or treating respiratory
PT   disorders e.g., lung cancer, chronic obstructive pulmonary disease,
PT   emphysema or asthma.
XX
PS   Claim 1; Page; 39pp; English.
XX
CC   The invention relates to a combination comprising cDNAs or their
CC   complements that are differentially expressed in respiratory disorder.
CC   The combination is useful for preparing a composition for diagnosing or
CC   treating respiratory disorders e.g. lung cancer, chronic obstructive
CC   pulmonary disease, emphysema or asthma. The present sequence represents
CC   human cDNA differentially expressed during lung cancer
XX
SQ   Sequence 1158 BP; 203 A; 413 C; 332 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.2e-129      Length:        1158
Score:                  1512.00       Matches:       275
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     9             Gaps:          0

US-09-598-982C-52 (1-275) x ACH04035 (1-1158)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db            798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 9
ADP10384
ID   ADP10384 standard; DNA; 1158 BP.
XX
AC   ADP10384;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Reference mRNA sequences for marker probe #61.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-00325899.
XX
PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J,  Fry K,  Woodward R,  Ly N,  Prentice J,  Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 80; SEQ ID NO 393; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC   and monitoring of allograft rejection and other disorders.
XX
SQ   Sequence 1158 BP; 203 A; 413 C; 332 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.2e-129       Length:         1158
Score:                  1512.00        Matches:        275
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     12             Gaps:           0

US-09-598-982C-52 (1-275) x ADP10384 (1-1158)

Qy              1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy             21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy             41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy             61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy             81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy            101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy            121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy            141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy            161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy            181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy            201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy            221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy            241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy            261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
                  |||||||||||||||||||||||||||||||||||||||||||||
Db            798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 10
ADP56069
ID   ADP56069 standard; cDNA; 1158 BP.
XX
AC   ADP56069;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human PRO cDNA sequence SEQ ID NO:2045.
XX
KW   human; PRO; immune related disease; inflammatory immune response;
KW   immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW   antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW   antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW   haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW   nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW   virucide; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004039956-A2.
XX
PD   13-MAY-2004.
```

```
XX
PF   28-OCT-2003; 2003WO-US034381.
XX
PR   29-OCT-2002; 2002US-0422472P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Aggarwal S,  Clark H,  Gurney AL,  Schoenfeld J,  Williams PM;
PI   Wood WI,  Wu TD;
XX
DR   WPI; 2004-376182/35.
DR   P-PSDB; ADP56070.
XX
PT   New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT   and treating an immune related disease, e.g. systemic lupus
PT   erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT   stimulating an immune response.
XX
PS   Claim 2; SEQ ID NO 2045; 3009pp; English.
XX
CC   The present invention describes an isolated PRO nucleic acid (I). Also
CC   described: (1) a vector comprising (I); (2) a host cell comprising the
CC   vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC   isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC   polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC   antibody which specifically binds to a polypeptide of (4); (7) a
CC   composition of matter comprising a polypeptide of (4), an agonist or
CC   antagonist of the polypeptide or an antibody that binds to the
CC   polypeptide in combination with a carrier; (8) an article of manufacture
CC   comprising a container, a label on the container and a composition of
CC   matter of (7); (9) a method of treating an immune related disease in a
CC   mammal; (10) a method for determining the presence of a PRO polypeptide
CC   in a sample suspected of having the polypeptide; (11) a method of
CC   diagnosing an immune related disease or an inflammatory immune response
CC   in mammal; (12) a method of identifying a compound that inhibits or
CC   mimics the activity of or expression of a gene encoding a PRO polypeptide
CC   ; and (13) a method of stimulating the immune response in a mammal. The
CC   PRO sequences have antiallergic, antianaemic, antiarthritic,
CC   antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC   antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC   haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC   nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC   virucide activities, and can be used in gene therapy. The nucleic acid
CC   (I) and the encoded polypeptides, compositions, kits and methods are
CC   useful in diagnosing and treating an immune related disease and in
CC   stimulating an immune response. The present sequence represents a human
CC   PRO nucleotide sequence from the present invention.
XX
SQ   Sequence 1158 BP; 203 A; 413 C; 332 G; 210 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:             5.2e-129      Length:        1158
Score:                 1512.00       Matches:       275
Percent Similarity:    100.00%       Conservative:  0
Best Local Similarity: 100.00%       Mismatches:    0
Query Match:           100.00%       Indels:        0
DB:                    13            Gaps:          0

US-09-598-982C-52 (1-275) x ADP56069 (1-1158)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 11
AAA34962
ID   AAA34962 standard; DNA; 17133 BP.
XX
AC   AAA34962;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human adenosine receptor related polynucleotide SEQ ID NO:2651.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US017712.
XX
PR   03-AUG-1998;   98US-0095212P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
```

```
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstruction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers.
XX
PS   Disclosure; Page 810-814; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC   sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC   from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
CC   AAA33992) are specifically claimed ONs from the present invention. N.B.
CC   Sequences given in the disclosure of the present invention do not match
CC   up with their corresponding SEQ ID NO: sequences given in the sequence
CC   listing
XX
SQ   Sequence 17133 BP; 2890 A; 5977 C; 5156 G; 3101 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:              1.54e-127     Length:        17133
Score:                  1512.00       Matches:       275
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     3             Gaps:          0

US-09-598-982C-52 (1-275) x AAA34962 (1-17133)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12636 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 12695

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12696 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 12755

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12756 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 12815

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12816 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 12875

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12876 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 12935

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12936 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 12995

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12996 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 13055

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13056 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 13115

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13116 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 13175

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13176 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 13235

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13236 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 13295

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13296 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 13355

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13356 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 13415

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db      13416 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 13460

RESULT 12
AAF21084
ID   AAF21084 standard; DNA; 17133 BP.
XX
AC   AAF21084;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Human low adenosine antisense oligonucleotide related sequence #2651.
XX
KW   Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW   human; airway disorder; bronchoconstriction; lung inflammation;
KW   surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW   immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW   respiratory obstruction; pulmonary obstruction; impeded respiration;
KW   surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW   respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW   pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW   chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200062736-A2.
XX
PD   26-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US008020.
XX
PR   06-APR-1999;   99US-0127958P.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
PA   (NYCE/) NYCE J W.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-679539/66.
XX
PT   Low adenosine (A) content antisense oligonucleotides which do not trigger
PT   adenosine receptors during metabolism, useful e.g. for treating cancers
```

```
PT   and respiratory obstructions.
XX
PS   Disclosure; Page 882-887; 1592pp; English.
XX
CC   The present invention describes low adenosine (A) content antisense
CC   oligonucleotides and compositions (I) comprising them. In the antisense
CC   oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC   (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC   immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC   The antisense oligonucleotides and (I) can be used to down-regulate the
CC   expression and or activity of target polypeptides associated with
CC   lung/respiratory disorders and malignancies, such as stimulating and
CC   activating peptide factors and transmitters, transcription factors,
CC   immunoglobulins and antibodies, antibody receptors, cytokines and
CC   chemokines, endogenously produced specific and non-specific enzymes,
CC   binding proteins, adhesion molecules and their receptors, cytokine and
CC   chemokine receptors, adenosine receptors, bradykinin receptors, central
CC   nervous system (CNS) and peripheral nervous and non-nervous system
CC   receptors, CNS and peripheral nervous and non-nervous system peptide
CC   transmitters, defensins, growth factors, vasoactive peptides and
CC   receptors, binding proteins and malignancy associated proteins. The
CC   antisense oligonucleotides may be used in this way to treat disorders
CC   including respiratory obstruction (especially pulmonary obstruction
CC   and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC   surfactant hypoproduction which are associated with a disease or
CC   condition selected from pulmonary vasoconstriction, inflammation,
CC   allergies, asthma, impeded respiration, respiratory distress syndrome
CC   (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC   pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC   and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC   fragments and antisense oligonucleotides used in the exemplification of
CC   the present invention
XX
SQ   Sequence 17133 BP; 2890 A; 5977 C; 5156 G; 3101 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:              1.54e-127      Length:        17133
Score:                  1512.00        Matches:       275
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     3              Gaps:          0

US-09-598-982C-52 (1-275) x AAF21084 (1-17133)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12636 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 12695

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12696 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 12755

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12756 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 12815

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12816 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 12875

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12876 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 12935

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12936 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 12995

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12996 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 13055
```

```
Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13056 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 13115

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13116 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 13175

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13176 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 13235

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13236 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 13295

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13296 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 13355

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13356 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 13415

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db      13416 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 13460

RESULT 13
ABZ96778
ID   ABZ96778 standard; DNA; 17133 BP.
XX
AC   ABZ96778;
XX
DT   17-OCT-2003  (first entry)
XX
DE   Human nucleic acid sequence.
XX
KW   Human; antisense; lung dysfunction; nasal airway dysfunction;
KW   antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW   antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW   antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW   adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW   lung inflammation; respiratory disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285308-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013135.
XX
PR   24-APR-2001; 2001US-0286137P.
XX
PA   (EPIG-) EPIGENESIS PHARM INC.
XX
PI   Nyce JW,  Li Y,  Sandrasagra A,  Katz E,  Pabalan J,  Aguilar D;
PI   Miller S,  Tang L,  Shahabuddin S;
XX
DR   WPI; 2003-229219/22.
XX
PT   Pharmaceutical composition for treating ailments associated with impaired
PT   respiration, has oligo(s) antisense to specific gene(s) or its
PT   corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT   ubiquinone.
XX
PS   Disclosure; SEQ ID NO 12020; 872pp; English.
XX
CC   The invention relates to a novel pharmaceutical composition, which has a
CC   first active agent comprising an oligonucleotide antisense to the
CC   initiation codon, coding region, 5' or 3' end genomic flanking regions,
```

```
CC   5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC   junctions of genes encoding a polypeptide associated with lung and/or
CC   nasal airway dysfunction and a second active agent comprising an
CC   antiinflammatory steroid and ubiquinone. A composition of the invention
CC   has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
CC   immunosuppressive, and cytostatic activity. The composition may have a
CC   use in antisense gene therapy. The composition is useful for treating or
CC   preventing a respiratory, lung or malignant disease or condition, also
CC   for enhancing the prophylactic or therapeutic respiratory effect of an
CC   antiinflammatory steroid in a subject, for reducing or depleting levels
CC   of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC   receptor, producing bronchodilation, increasing levels of ubiquinone or
CC   lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC   lung inflammation, lung allergies, or a respiratory disease or condition.
CC   Note: The sequence data for this patent is not represented in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 17133 BP; 2890 A; 5977 C; 5156 G; 3101 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:              1.54e-127      Length:         17133
Score:                  1512.00        Matches:        275
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     10             Gaps:           0

US-09-598-982C-52 (1-275) x ABZ96778 (1-17133)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12636 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 12695

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12696 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 12755

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12756 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 12815

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12816 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 12875

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12876 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 12935

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12936 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 12995

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12996 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 13055

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13056 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 13115

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13116 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 13175

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13176 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 13235

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13236 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 13295
```

```
Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13296 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 13355

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13356 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 13415

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db      13416 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 13460

RESULT 14
ABD20627
ID   ABD20627 standard; DNA; 17133 BP.
XX
AC   ABD20627;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human pulmonary and inflammatory target DNA #238.
XX
KW   Human; antisense; bronchoconstriction; allergy; hyposecretion; pain;
KW   respiratory tract inflammation; adenosine sensitivity; lung; cancer;
KW   surfactant depletion; antiallergic; antiinflammatory; antiasthmatic;
KW   analgesic; hypotensive; immunosuppressive; cytostatic; cystic fibrosis;
KW   beta-adrenergic agonist; respiratory disease; pulmonary vasoconstriction;
KW   respiratory distress syndrome; allergic rhinitis; pulmonary hypertension;
KW   emphysema; chronic obstructive pulmonary disease; cancer; bronchitis;
KW   pulmonary transplantation rejection; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285309-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013143.
XX
PR   24-APR-2001; 2001US-0286036P.
XX
PA   (EPIG-) EPIGENESIS PHARM INC.
XX
PI   Nyce JW,  Li Y,  Sandrasagra A,  Katz E,  Pabalan J,  Aguilar D;
PI   Miller S,  Tang L,  Shahabuddin S;
XX
DR   WPI; 2003-093058/08.
XX
PT   Pharmaceutical composition for treating asthma, has antisense
PT   oligonucleotide containing less percentage of adenosine, targeted to
PT   nucleic acids associated with lung airway or lung dysfunction, and
PT   bronchodilating agent.
XX
PS   Claim 15; SEQ ID NO 12020; 763pp; English.
XX
CC   This invention describes a novel composition (a) a first active agent,
CC   comprising oligonucleotides, effective for alleviating
CC   bronchoconstriction, respiratory tract inflammation, allergies and
CC   reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors,
CC   surfactant depletion or hyposecretion, when administered to a mammal. The
CC   oligonucleotides are derived from a gene encoding or regulating
CC   expression of a target polypeptide associated with lung airway or lung
CC   dysfunction or cancer and can be anti-sense to the corresponding mRNA.
CC   The invention also describes a kit, that comprises: (a) a delivery
CC   device, in separate containers, (b) the oligonucleotides, (c)
CC   instructions for adding a carrier and for use of the kit. The composition
CC   of the invention has antiallergic, antiinflammatory, antiasthmatic,
CC   analgesic, hypotensive, immunosuppressive and cytostatic activity, is a
CC   beta-adrenergic agonist. The composition is useful for preventing or
CC   treating a respiratory, lung or malignant disease. The administered
CC   composition comprises oligo and is administered to reduce the production
CC   or availability, or to increase the degradation of the target mRNA or to
```

```
CC   reduce the amount of target polypeptide present in the lungs. The
CC   pulmonary obstruction, and/or bronchoconstriction and/or lung
CC   inflammation, allergies and/or surfactant hypoproduction are associated
CC   with a disease or condition such as pulmonary vasoconstriction,
CC   inflammation, allergies, asthma, impeded respiration, respiratory
CC   distress syndrome, pain, cystic fibrosis, allergic rhinitis, pulmonary
CC   hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary
CC   transplantation rejection, pulmonary infections, bronchitis or cancer.
CC   The reduced adenosine content of the anti-sense oligos corresponding to
CC   thymidines present in the target RNA serves to prevent the breakdown of
CC   the oligonucleotides into products that free adenosine into the system
CC   e.g., lung, brain, heart, kidney, etc, tissue environment and thereby, to
CC   prevent any unwanted effects due to it
XX
SQ   Sequence 17133 BP; 2890 A; 5977 C; 5156 G; 3101 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:              1.54e-127     Length:        17133
Score:                  1512.00       Matches:       275
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     11            Gaps:          0

US-09-598-982C-52 (1-275) x ABD20627 (1-17133)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12636 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 12695

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12696 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 12755

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12756 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 12815

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12816 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 12875

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12876 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 12935

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12936 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 12995

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12996 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 13055

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13056 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 13115

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13116 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 13175

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13176 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 13235

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13236 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 13295

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13296 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 13355

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13356 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 13415

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db      13416 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 13460

RESULT 15
ADQ85402
ID   ADQ85402 standard; cDNA; 1142 BP.
XX
AC   ADQ85402;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human tumour-associated antigenic target (TAT) cDNA sequence #2216.
XX
KW   human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW   cancer; cell proliferative disorder; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004060270-A2.
XX
PD   22-JUL-2004.
XX
PF   15-OCT-2003; 2003WO-US029126.
XX
PR   18-OCT-2002; 2002US-0418988P.
XX
PA   (GETH ) GENENTECH INC.
PA   (WUTD/) WU T D.
PA   (ZHOU/) ZHOU Y.
XX
PI   Wu TD,  Zhou Y;
XX
DR   WPI; 2004-534300/51.
XX
PT   New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT   preventing or treating cell proliferative disorders such as cancer.
XX
PS   Claim 1; SEQ ID NO 2216; 5504pp; English.
XX
CC   The present invention describes an isolated tumour-associated antigenic
CC   target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC   sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC   (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC   sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC   (c). Also described: (1) an expression vector comprising the above
CC   nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC   a process for producing a polypeptide; (4) an isolated polypeptide
CC   comprising: (a) an amino acid sequence encoded by any of the above
CC   nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC   length coding region of the above nucleotide sequences; or (c) a sequence
CC   having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC   comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC   an isolated antibody that binds to the above polypeptide; (7) a process
CC   for producing the antibody; (8) an isolated oligopeptide that binds to
CC   the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC   binding organic molecule that binds to the above polypeptide; (10) a
CC   composition of matter comprising the above (chimeric) polypeptide,
CC   antibody, oligopeptide or TAT binding organic molecule, in combination
CC   with a carrier; (11) an article of manufacture comprising a container and
CC   the composition of matter contained within the container; (12) methods of
CC   inhibiting the growth of a cell that expresses the above protein, where
CC   the growth of the cell is at least in part dependent upon a growth
CC   potentiating effect of the above protein; (13) a method of
CC   therapeutically treating a mammal having a cancerous tumour comprising
CC   cells that express the above protein; (14) a method of determining the
CC   presence of a protein in a sample suspected of containing the protein
CC   described above; (15) methods of diagnosing the presence of a tumour in a
```

```
CC   mammal; (16) a method for treating or preventing a cell proliferative
CC   disorder associated with increased expression or activity of the above
CC   protein; and (17) a method of binding an antibody, oligopeptide or
CC   organic molecule to a cell that expresses the protein described above.
CC   The TAT sequences have cytostatic activities, and can be used in gene
CC   therapy. The composition and methods are useful for diagnosing,
CC   preventing or treating cancer. The composition is also used for preparing
CC   a medicament for the therapeutic treatment or diagnostic detection of a
CC   cell proliferative disorder or cancer. The present sequence represents a
CC   human TAT cDNA sequence from the present invention.
XX
SQ   Sequence 1142 BP; 194 A; 412 C; 333 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.69e-121     Length:        1142
Score:                  1430.00       Matches:       257
Percent Similarity:     96.36%        Conservative:  8
Best Local Similarity:  93.45%        Mismatches:    10
Query Match:            94.58%        Indels:        0
DB:                     13            Gaps:          0

US-09-598-982C-52 (1-275) x ADQ85402 (1-1142)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 61

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         62 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 121

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 181

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        182 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTC 241

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        242 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 301

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        302 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 361

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        362 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 421

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 481

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        482 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 541

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 601

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        602 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 661

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 721

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        722 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 781

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        782 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 826

Search completed: August 30, 2005, 20:29:19
Job time : 6436 secs
```

```
                        GenCore version 5.1.6
                 Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         August 30, 2005, 15:02:36 ; Search time 3552 Seconds
                                          (without alignments)
                                          3751.460 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters:      9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_29082005_080946_26113/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_5600_@runat_29082005_080946_26113 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                           Description
-------------------------------------------------------------------------------
    1    1512   100.0   1154  6  AR080459                     AR080459 Sequence
    2    1512   100.0   1154  9  HUMTRY                       M30038 Human trypt
    3    1494    98.8    884  9  AF206665                     AF206665 Homo sapi
    4    1494    98.8   1206  9  BC028059                     BC028059 Homo sapi
    5    1494    98.8   1207  9  BC051852                     BC051852 Homo sapi
    6    1414    93.5    896  9  BC074975                     BC074975 Homo sapi
    7    1405    92.9    896  9  BC074974                     BC074974 Homo sapi
    8    1405    92.9   1143  9  HUMBTRYP                     M37488 Human beta-
    9    1405    92.9   1145  9  S55551                       S55551 beta-trypta
   10    1405    92.9   1194  9  BC029356                     BC029356 Homo sapi
   11    1405    92.9   1235  9  BC038416                     BC038416 Homo sapi
   12    1405    92.9   1460  9  BC065923                     BC065923 Homo sapi
   13    1403    92.8    825  9  CR541836                     CR541836 Homo sapi
   14    1402    92.7   1137  6  AR080460                     AR080460 Sequence
   15    1402    92.7   1137  9  HUMTRY1A                     M33491 Human trypt
   16    1401    92.7   2662  6  AX338508                     AX338508 Sequence
   17    1400    92.6   1128  6  AR080461                     AR080461 Sequence
   18    1400    92.6   1128  9  HUMTRY2A                     M33492 Human trypt
   19    1394    92.2   1081  6  AR080462                     AR080462 Sequence
   20    1394    92.2   1081  6  AR482491                     AR482491 Sequence
   21    1394    92.2   1081  6  AX329646                     AX329646 Sequence
   22    1394    92.2   1081  6  AX330205                     AX330205 Sequence
   23    1394    92.2   1081  6  AX334644                     AX334644 Sequence
   24    1394    92.2   1081  6  AX375047                     AX375047 Sequence
   25    1394    92.2   1081  9  HUMTRY3A                     M33493 Human trypt
   26  1378.5    91.2    768  9  AF206666                     AF206666 Homo sapi
   27    1364    90.2   1221  6  BD222223                     BD222223 Human nuc
   28    1364    90.2   1221  6  AX014344                     AX014344 Sequence
   29    1348    89.2   5441  9  AF098328                     AF098328 Homo sapi
   30    1290    85.3    735  6  AX347834                     AX347834 Sequence
   31    1290    85.3    771  6  AX347838                     AX347838 Sequence
   32    1287    85.1    735  6  AX347881                     AX347881 Sequence
   33    1287    85.1    735  6  AX347883                     AX347883 Sequence
   34    1287    85.1    771  6  AX347873                     AX347873 Sequence
   35    1287    85.1    771  6  AX347875                     AX347875 Sequence
   36    1284    84.9    735  6  AR165109                     AR165109 Sequence
   37    1284    84.9    735  6  BD211487                     BD211487 Enzymatic
   38    1284    84.9    735  6  AX347843                     AX347843 Sequence
   39    1284    84.9    771  6  AR165112                     AR165112 Sequence
   40    1284    84.9    771  6  BD211490                     BD211490 Enzymatic
   41    1284    84.9    771  6  AX347841                     AX347841 Sequence
   42    1282    84.8    735  6  AX347879                     AX347879 Sequence
   43    1282    84.8    771  6  AX347871                     AX347871 Sequence
   44    1281    84.7    735  6  AX347865                     AX347865 Sequence
   45    1281    84.7    735  6  AX347867                     AX347867 Sequence

                                 ALIGNMENTS

RESULT 1
AR080459
LOCUS       AR080459                1154 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 11 from patent US 5968782.
ACCESSION   AR080459
VERSION     AR080459.1  GI:10007194
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1154)
  AUTHORS   Stevens,R.L.
  TITLE     Mast cell protease that cleaves fibrinogen
  JOURNAL   Patent: US 5968782-A 11 19-OCT-1999;
FEATURES             Location/Qualifiers
     source          1..1154
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              7.85e-136      Length:          1154
Score:                  1512.00        Matches:         275
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     6              Gaps:            0
```

```
US-09-598-982C-52 (1-275) x AR080459 (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 2
HUMTRY
LOCUS       HUMTRY                  1154 bp    mRNA    linear   PRI 07-FEB-1996
DEFINITION  Human tryptase mRNA, complete cds.
ACCESSION   M30038
VERSION     M30038.1  GI:1182066
KEYWORDS    serine protease; tryptase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1154)
  AUTHORS   Miller,J.S., Westin,E.H. and Schwartz,L.B.
  TITLE     Cloning and characterization of complementary DNA for human
            tryptase
  JOURNAL   J. Clin. Invest. 84 (4), 1188-1195 (1989)
  MEDLINE   90009311
   PUBMED   2677049
REFERENCE   2  (bases 1 to 1154)
  AUTHORS   Schwartz,L.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAR-1990) Lawrence B. Schwartz, Medicine, Virginia
            Commonwealth University, Richmond, VA 23298, USA
COMMENT     On Feb 7, 1996 this sequence version replaced gi:339978.
            Original source text: Human cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .1154
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="lung"
     mRNA            <1. .1154
                     /product="TRY mRNA"
     CDS             18. .845
                     /codon_start=1
                     /product="tryptase precursor"
                     /protein_id="AAA86934.1"
                     /db_xref="GI:1182067"
                     /translation="MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPW
                     QVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGW
                     GDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRD
                     SCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
     sig_peptide     18. .107
     mat_peptide     108. .839
                     /product="tryptase"
ORIGIN

Alignment Scores:
Pred. No.:              7.85e-136       Length:         1154
Score:                  1512.00         Matches:        275
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-598-982C-52 (1-275) x HUMTRY (1-1154)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 77

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 137

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 197

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 257

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 317

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CTGCTGCCAGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 377

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 437

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        438 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 497

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 557

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 617

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGC 677

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGGGCGACTCTGGAGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 737

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 797

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        798 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 842

RESULT 3
AF206665
LOCUS       AF206665                 884 bp    mRNA    linear   PRI 24-OCT-2002
DEFINITION  Homo sapiens mast cell alpha II tryptase mRNA, complete cds,
            alternatively spliced.
ACCESSION   AF206665
VERSION     AF206665.1  GI:11493897
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   Wang,H.W., McNeil,H.P., Thomas,P.S., Murphy,B.N., Webster,M.J.,
            Hettiaratchi,A., King,G., Heywood,G.J., Huang,C., Stevens,R.L. and
            Hunt,J.E.
  TITLE     delta Tryptase Is Expressed in Multiple Human Tissues, and a
            Recombinant Form Has Proteolytic Activity
  JOURNAL   J. Immunol. 169 (9), 5145-5152 (2002)
  MEDLINE   22278650
   PUBMED   12391231
REFERENCE   2  (bases 1 to 884)
  AUTHORS   Hunt,J.E., Wang,H.W., Thomas,H.P. and McNeil,H.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-1999) Pathology, University of New South Wales,
            High St., Kensington, NSW 2052, Australia
FEATURES             Location/Qualifiers
     source          1..884
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="16"
                     /map="16p13.3"
                     /tissue_type="lung"
                     /clone_lib="Invitrogen cat. number D6090-15"
     CDS             11..838
                     /note="serine protease; alternatively spliced; similar to
                     the product encoded by GenBank Accession Number AF098328"
                     /codon_start=1
                     /product="mast cell alpha II tryptase"
                     /protein_id="AAG35695.1"
                     /db_xref="GI:11493898"
                     /translation="MLSLLLLALPVLASPAYAAPAPVQALQQAGIVGGQEAPRSKWPW
                     QVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGW
                     GDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
```

```
ORIGIN

Alignment Scores:
Pred. No.:              3.11e-134      Length:          884
Score:                  1494.00        Matches:         271
Percent Similarity:     99.64%         Conservative:    3
Best Local Similarity:  98.55%         Mismatches:      1
Query Match:            98.81%         Indels:          0
DB:                     9              Gaps:            0

US-09-598-982C-52 (1-275) x AF206665 (1-884)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              |||||||||||||||||||||||||||||||||||||||      |||||||||||||||
Db         11 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGCGGCCCCT 70

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 130

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGTGGG 190

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 250

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 310

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 370

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 430

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 490

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 550

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 610

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||::::::||||||||||||
Db        611 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 670

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 CAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 730

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 790

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        791 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 835

RESULT 4
BC028059
LOCUS       BC028059                1206 bp    mRNA    linear   PRI 06-OCT-2003
DEFINITION  Homo sapiens tryptase, alpha, mRNA (cDNA clone MGC:39869
            IMAGE:5181815), complete cds.
```

```
ACCESSION   BC028059
VERSION     BC028059.1  GI:20380844
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1206)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1206)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 62 Row: e Column: 8
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 13699841.
FEATURES             Location/Qualifiers
     source          1. .1206
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:39869 IMAGE:5181815"
                     /tissue_type="Brain, Lung, Testis, adult, pooled whole"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1206
                     /gene="TPS1"
                     /note="synonyms: TPS2, alpha II"
                     /db_xref="LocusID:7176"
                     /db_xref="MIM:191080"
     CDS             13. .840
                     /gene="TPS1"
                     /codon_start=1
                     /product="TPS1 protein"
                     /protein_id="AAH28059.1"
                     /db_xref="GI:20380845"
                     /db_xref="LocusID:7176"
                     /db_xref="MIM:191080"
                     /translation="MLSLLLLALPVLASPAYAAPAPVQALQQAGIVGGQEAPRSKWPW
                     QVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGW
                     GDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
     misc_feature    103. .813
                     /gene="TPS1"
                     /note="Tryp_SPc; Region: Trypsin-like serine protease"
                     /db_xref="CDD:smart00020"
ORIGIN

Alignment Scores:
Pred. No.:              4.46e-134      Length:         1206
Score:                  1494.00        Matches:        271
Percent Similarity:     99.64%         Conservative:   3
Best Local Similarity:  98.55%         Mismatches:     1
Query Match:            98.81%         Indels:         0
DB:                     9              Gaps:           0

US-09-598-982C-52 (1-275) x BC028059 (1-1206)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db         13 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGCGGCCCCT 72

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 132

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGTGGG 192

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 252

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 312

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 372

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 432

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 492

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 552

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 612
```

```
Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||::::::||||||||||||
Db        613 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 672

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 CAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 732

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 792

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        793 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 837

RESULT 5
BC051852
LOCUS       BC051852                1207 bp    mRNA    linear   PRI 17-JUN-2003
DEFINITION  Homo sapiens tryptase, alpha, mRNA (cDNA clone IMAGE:5181454),
            partial cds.
ACCESSION   BC051852
VERSION     BC051852.1  GI:30704872
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1207)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE  22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1207)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.
```

```
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 110 Row: b Column: 1
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 13699841.
FEATURES             Location/Qualifiers
     source          1..1207
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5181454"
                     /tissue_type="Brain, Lung, Testis, adult, pooled whole"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            <1..1207
                     /gene="TPS1"
                     /note="synonyms: TPS2, alpha II"
                     /db_xref="LocusID:7176"
                     /db_xref="MIM:191080"
     CDS             <1..833
                     /gene="TPS1"
                     /codon_start=3
                     /product="TPS1 protein"
                     /protein_id="AAH51852.1"
                     /db_xref="GI:30704873"
                     /db_xref="LocusID:7176"
                     /db_xref="MIM:191080"
                     /translation="RMLSLLLLALPVLASPAYAAPAPVQALQQAGIVGGQEAPRSKWP
                     WQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQL
                     LPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTG
                     WGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRR
                     DSCQGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP
                     "
ORIGIN

Alignment Scores:
Pred. No.:              4.47e-134     Length:         1207
Score:                  1494.00       Matches:        271
Percent Similarity:     99.64%        Conservative:   3
Best Local Similarity:  98.55%        Mismatches:     1
Query Match:            98.81%        Indels:         0
DB:                     9             Gaps:           0

US-09-598-982C-52 (1-275) x BC051852 (1-1207)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db          6 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGCGGCCCCT 65

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 125

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGTGGG 185

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 245

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 305

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 365

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 425
```

```
Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 485

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 545

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 605

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              |||||||||||||||||||||||||||||||||||||||::::::|||||||||||||||
Db        606 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 665

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 725

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTGGTCAGCTGGGACGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 785

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        786 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 830

RESULT 6
BC074975
LOCUS       BC074975                 896 bp    mRNA    linear   PRI 04-AUG-2004
DEFINITION  Homo sapiens tryptase beta 1, mRNA (cDNA clone MGC:103905
            IMAGE:30915294), complete cds.
ACCESSION   BC074975
VERSION     BC074975.2  GI:50960667
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 896)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Aug 4, 2004 this sequence version replaced gi:49901651.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Genome Sequence Centre, British Columbia Cancer
            Center
            cDNA Library Preparation: British Columbia Cancer Research Center
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRBU Plate: 2 Row: C Column: 7.
FEATURES             Location/Qualifiers
     source          1..896
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:103905 IMAGE:30915294"
                     /tissue_type="Pooled tissue, PCR rescued clones"
                     /clone_lib="NIH_MGC_271"
                     /note="Vector: pCR4 Topo TA with reversed insert"
     gene            1..896
                     /gene="TPSB1"
                     /note="synonyms: TPS2, beta I, tryptase-II, TPS1"
                     /db_xref="LocusID:7177"
                     /db_xref="MIM:191081"
     CDS             34..861
                     /gene="TPSB1"
                     /codon_start=1
                     /product="beta tryptase 1, precursor"
                     /protein_id="AAH74975.1"
                     /db_xref="GI:50960668"
                     /db_xref="LocusID:7177"
                     /db_xref="MIM:191081"
                     /translation="MLSLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPW
                     QVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGW
                     GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.58e-126       Length:         896
Score:                  1414.00         Matches:        255
Percent Similarity:     95.27%          Conservative:   7
Best Local Similarity:  92.73%          Mismatches:     13
Query Match:            93.52%          Indels:         0
DB:                     9               Gaps:           0

US-09-598-982C-52 (1-275) x BC074975 (1-896)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 93

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   ||||||||||||||||||||||||||||||
Db         94 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 153

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        154 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 213

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
```

```
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        214 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTC 273

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        274 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 333

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        334 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 393

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        394 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 453

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 513

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        514 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 573

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 633

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        634 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 693

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 753

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        754 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 813

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        814 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 858

RESULT 7
BC074974
LOCUS       BC074974                 896 bp    mRNA    linear   PRI 04-AUG-2004
DEFINITION  Homo sapiens tryptase beta 2, mRNA (cDNA clone MGC:104064
            IMAGE:30915512), complete cds.
ACCESSION   BC074974
VERSION     BC074974.2  GI:50959860
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 896)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Aug 4, 2004 this sequence version replaced gi:49901868.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Genome Sequence Centre, British Columbia Cancer
            Center
            cDNA Library Preparation: British Columbia Cancer Research Center
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRBU Plate: 4 Row: E Column: 9.
FEATURES             Location/Qualifiers
     source          1. .896
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:104064 IMAGE:30915512"
                     /tissue_type="Pooled tissue, PCR rescued clones"
                     /clone_lib="NIH_MGC_271"
                     /note="Vector: pCR4 Topo TA with reversed insert"
     gene            1. .896
                     /gene="TPSB2"
                     /db_xref="LocusID:64499"
     CDS             34. .861
                     /gene="TPSB2"
                     /codon_start=1
                     /product="tryptase beta 2, precursor"
                     /protein_id="AAH74974.1"
                     /db_xref="GI:50959861"
                     /db_xref="LocusID:64499"
                     /translation="MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPW
                     QVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGW
                     GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.16e-125      Length:         896
Score:                  1405.00        Matches:        253
Percent Similarity:     94.91%         Conservative:   8
Best Local Similarity:  92.00%         Mismatches:     14
Query Match:            92.92%         Indels:         0
DB:                     9              Gaps:           0

US-09-598-982C-52 (1-275) x BC074974 (1-896)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         34 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 93

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         94 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 153

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        154 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 213

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        214 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 273

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        274 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 333

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        334 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 393

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        394 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 453

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 513

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        514 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 573

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 633

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        634 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 693

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 753

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        754 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 813

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        814 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 858

RESULT 8
HUMBTRYP
LOCUS       HUMBTRYP                1143 bp    mRNA    linear   PRI 31-OCT-1994
DEFINITION  Human beta-tryptase mRNA, complete cds.
ACCESSION   M37488
VERSION     M37488.1  GI:179583
KEYWORDS    beta-tryptase; mast cell serine protease; tryptase.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1143)
  AUTHORS   Miller,J.S., Moxley,G. and Schwartz,L.B.
  TITLE     Cloning and characterization of a second complementary DNA for
            human tryptase
  JOURNAL   J. Clin. Invest. 86 (3), 864-870 (1990)
  MEDLINE   90369005
   PUBMED   2203827
COMMENT     Original source text: Human adult lung mast cell, cDNA to mRNA,
            clone 10.
            Draft entry and computer-readable sequence for [J. Clin. Invest.
            (1990) In press] kindly submitted
            by L.B.Schwartz, 06-AUG-1990.
FEATURES             Location/Qualifiers
     source          1. .1143
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="16"
                     /cell_type="mast cell"
                     /tissue_type="lung"
     gene            1. .1143
                     /gene="TPS2"
     mRNA            <1. .1143
                     /gene="TPS2"
                     /product="beta-tryptase"
     CDS             6. .833
                     /gene="TPS2"
                     /codon_start=1
                     /product="beta-tryptase"
                     /protein_id="AAA51843.1"
                     /db_xref="GI:179584"
                     /db_xref="GDB:G00-125-891"
                     /translation="MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPW
                     QVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGW
                     GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
     sig_peptide     6. .95
                     /gene="TPS2"
                     /note="G00-125-891"
     mat_peptide     96. .830
                     /gene="TPS2"
                     /product="beta-tryptase"
                     /note="G00-125-891"
ORIGIN

Alignment Scores:
Pred. No.:              1.54e-125       Length:         1143
Score:                  1405.00         Matches:        253
Percent Similarity:     94.91%          Conservative:   8
Best Local Similarity:  92.00%          Mismatches:     14
Query Match:            92.92%          Indels:         0
DB:                     9               Gaps:           0

US-09-598-982C-52 (1-275) x HUMBTRYP (1-1143)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 65

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         66 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 125

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        126 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 185

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        186 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 245

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        246 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 305

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        306 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 365

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
```

```
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        366 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 425

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 485

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        486 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 545

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 605

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        606 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 665

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 725

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 785

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        786 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 830

RESULT 9
S55551
LOCUS       S55551                  1145 bp    mRNA    linear   PRI 08-MAY-1993
DEFINITION  beta-tryptase [human, basophil cell line KU812, mRNA, 1145 nt].
ACCESSION   S55551
VERSION     S55551.1  GI:265666
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1145)
  AUTHORS   Blom,T. and Hellman,L.
  TITLE     Characterization of a tryptase mRNA expressed in the human basophil
            cell line KU812
  JOURNAL   Scand. J. Immunol. 37 (2), 203-208 (1993)
  MEDLINE   93166209
   PUBMED   8434231
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 125219] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .1145
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .1145
                     /gene="beta-tryptase"
                     /note="precursor"
     CDS             8. .835
                     /gene="beta-tryptase"
                     /codon_start=1
                     /protein_id="AAD13876.1"
                     /db_xref="GI:4261576"
                     /translation="MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPW
                     QVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLL
                     PVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGW
                     GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:             1.54e-125      Length:        1145
```

```
Score:                 1405.00        Matches:        253
Percent Similarity:    94.91%         Conservative:   8
Best Local Similarity: 92.00%         Mismatches:     14
Query Match:           92.92%         Indels:         0
DB:                    9              Gaps:           0

US-09-598-982C-52 (1-275) x S55551 (1-1145)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 67

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         68 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 127

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        128 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 187

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        188 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 247

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        248 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 307

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        308 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 367

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        368 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 427

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 487

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        488 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 547

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 607

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        608 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 667

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 727

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        728 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 787

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        788 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 832

RESULT 10
BC029356
LOCUS       BC029356                1194 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens tryptase beta 2, mRNA (cDNA clone MGC:32523
            IMAGE:4695458), complete cds.
ACCESSION   BC029356
VERSION     BC029356.1  GI:20809954
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
```

```
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 1194)
  AUTHORS  Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
           Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
           Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
           Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
           Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
           Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
           Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
           Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
           Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
           McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
           Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
           Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
           Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
           Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
           Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
           Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
           Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
           Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE    Generation and initial analysis of more than 15,000 full-length
           human and mouse cDNA sequences
  JOURNAL  Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED  12477932
REFERENCE  2  (bases 1 to 1194)
  AUTHORS  Strausberg,R.
  TITLE    Direct Submission
  JOURNAL  Submitted (01-MAY-2002) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT    Contact: MGC help desk
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: CLONTECH
           cDNA Library Preparation: CLONTECH Laboratories, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Sequencing Group at the Stanford Human Genome
           Center, Stanford University School of Medicine, Stanford, CA  94305
           Web site:       http://www-shgc.stanford.edu
           Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
           Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
           R. M.

           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAL Plate: 41 Row: j Column: 11
           This clone was selected for full length sequencing because it
           passed the following selection criteria: matched mRNA gi: 13775596.
FEATURES            Location/Qualifiers
     source         1..1194
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="MGC:32523 IMAGE:4695458"
                    /tissue_type="Lung"
                    /clone_lib="NIH_MGC_77"
                    /lab_host="DH10B"
                    /note="Vector: pDNR-LIB"
     gene           1..1194
                    /gene="TPSB2"
                    /db_xref="LocusID:64499"
     CDS            24..851
                    /gene="TPSB2"
                    /codon_start=1
                    /product="tryptase beta 2, precursor"
                    /protein_id="AAH29356.1"
                    /db_xref="GI:20809955"
                    /db_xref="LocusID:64499"
                    /translation="MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPW
                    QVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLL
                    PVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGW
                    GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRD
                    SCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.62e-125      Length:         1194
Score:                  1405.00        Matches:        253
Percent Similarity:     94.91%         Conservative:   8
Best Local Similarity:  92.00%         Mismatches:     14
Query Match:            92.92%         Indels:         0
DB:                     9              Gaps:           0

US-09-598-982C-52 (1-275) x BC029356 (1-1194)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 83

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         84 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 143

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        144 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 203

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        204 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 263

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        264 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 323

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        324 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 383

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        384 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 443

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 503

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        504 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 563

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 623

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        624 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 683

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 743

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        744 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 803

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        804 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 848

RESULT 11
BC038416
```

```
LOCUS       BC038416                1235 bp    mRNA    linear   PRI 25-JUN-2004
DEFINITION  Homo sapiens tryptase beta 2, mRNA (cDNA clone IMAGE:5176649),
            partial cds.
ACCESSION   BC038416
VERSION     BC038416.1  GI:23512304
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1235)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1235)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 51 Row: n Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 13775596.
FEATURES             Location/Qualifiers
     source          1..1235
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5176649"
                     /tissue_type="Brain, Lung, Testis, adult, pooled whole"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
```

```
Pred. No.:              1.69e-125     Length:         1235
Score:                  1405.00       Matches:        253
Percent Similarity:     94.91%        Conservative:   8
Best Local Similarity:  92.00%        Mismatches:     14
Query Match:            92.92%        Indels:         0
DB:                     9             Gaps:           0

US-09-598-982C-52 (1-275) x BC038416 (1-1235)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 66

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         67 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 126

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        127 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 186

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        187 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 246

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        247 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 306

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        307 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 366

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        367 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 426

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 486

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        487 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 546

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 606

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        607 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 666

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 726

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        727 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 786

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        787 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 831

RESULT 12
BC065923
LOCUS       BC065923                1460 bp    mRNA    linear   PRI 23-AUG-2004
DEFINITION  Homo sapiens tryptase beta 2, mRNA (cDNA clone MGC:74871
            IMAGE:5752632), complete cds.
ACCESSION   BC065923
VERSION     BC065923.1  GI:41944664
KEYWORDS    MGC.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1460)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1460)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 139 Row: n Column: 24
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 13775596.
FEATURES             Location/Qualifiers
     source          1. .1460
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:74871 IMAGE:5752632"
                     /tissue_type="Brain, Lung, Testis, adult, pooled whole"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1460
                     /gene="TPSB2"
                     /db_xref="LocusID:64499"
     CDS             191. .1039
```

```
                     /gene="TPSB2"
                     /codon_start=1
                     /product="TPSB2 protein"
                     /protein_id="AAH65923.1"
                     /db_xref="GI:41944665"
                     /db_xref="LocusID:64499"
                     /translation="MLLLAPQMLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEA
                     PRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHL
                     YYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGM
                     PCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLC
                     AGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHH
                     YVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-125       Length:         1460
Score:                  1405.00         Matches:        253
Percent Similarity:     94.91%          Conservative:   8
Best Local Similarity:  92.00%          Mismatches:     14
Query Match:            92.92%          Indels:         0
DB:                     9               Gaps:           0

US-09-598-982C-52 (1-275) x BC065923 (1-1460)

Qy           1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
               ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         212 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 271

Qy          21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
               ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         272 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 331

Qy          41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
               ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db         332 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 391

Qy          61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
               |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db         392 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 451

Qy          81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
               ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db         452 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 511

Qy         101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
               ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db         512 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 571

Qy         121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
               |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db         572 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 631

Qy         141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         632 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 691

Qy         161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         692 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 751

Qy         181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         752 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 811

Qy         201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
               ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db         812 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 871

Qy         221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
               :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         872 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 931

Qy         241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
```

```
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        932 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 991

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        992 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 1036

RESULT 13
CR541836
LOCUS       CR541836                 825 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens full open reading frame cDNA clone RZPDo834B0932D for
            gene TPSB2, tryptase beta 2; complete cds, without stopcodon.
ACCESSION   CR541836
VERSION     CR541836.1  GI:49456628
KEYWORDS    Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   Halleck,A., Ebert,L., Mkoundinya,M., Schick,M., Eisenstein,S.,
            Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,
            Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
  TITLE     Cloning of human full open reading frames in Gateway(TM) system
            entry vector (pDONR201)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 825)
  AUTHORS   Halleck,A., Ebert,L., Mkoundinya,M., Schick,M., Eisenstein,S.,
            Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,
            Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
            Germany
COMMENT     RZPD; RZPDo834B0932D, ORFNo 3770
            www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDo834B0932D RZPDLIB;
            Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
            834
            www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
            www.rzpd.de/products/orfclones/
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            Contact RZPD (customer.service@rzpd.de) for further information.
            Clone name at Harvard Institute of Proteomics
            (www.hip.harvard.edu): FLH130921.01L
            This CDS clone is part of a collection of human full ORF clones
            jointly established and verified by the Harvard Institute of
            Proteomics (HIP) and RZPD.
            This CDS has been cloned without stopcodon.
            The CDS has been inserted into pDONR201 via a BP Clonase(TM)
            reaction. Additional sequence has been added in front of the start
            codon: att. .AAAAAA GCA GGC TCC ACC (ATG).
            The last codon is followed by the 3' att site: GACCCAGCTTTCTT. .att
            The clone is validated by full sequence check.
            Compared to the reference sequence NM_003294 (GI:13775594) we found
            AA exchange(s) at position (first base of changed triplet):
            184(ser->thr) 319(ile->thr)
            Clone distribution: http://www.rzpd.de/products/orfclones/.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDo834B0932D"
                     /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
                     /lab_host="DH5Alpha"
                     /note="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
     gene            1. .825
                     /gene="TPSB2"
     CDS             1. .>825
                     /gene="TPSB2"
                     /codon_start=1
                     /protein_id="CAG46635.1"
                     /db_xref="GI:49456629"
                     /translation="MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPW
                     QVSLRVHGPYWMHFCGGTLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLL
                     PVSRTIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGW
                     GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRD
                     SCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.64e-125      Length:         825
Score:                  1403.00        Matches:        252
Percent Similarity:     94.91%         Conservative:   9
Best Local Similarity:  91.64%         Mismatches:     14
Query Match:            92.79%         Indels:         0
DB:                     9              Gaps:           0

US-09-598-982C-52 (1-275) x CR541836 (1-825)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 60

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         61 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 120

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        121 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 180

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||:::|||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        181 GGCACCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 240

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 300

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||   |||||||||||||||||||||      |||   ||||||
Db        301 CTGCTGCCGGTCAGCAGGACCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 360

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        361 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 420

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 480

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        481 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 540

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 600

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db        601 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 660

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGC 720

Qy        241 ValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgVal 260
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        721 GTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTC 780

Qy        261 ThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 825

RESULT 14
AR080460
LOCUS       AR080460                1137 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 13 from patent US 5968782.
ACCESSION   AR080460
VERSION     AR080460.1  GI:10007195
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1137)
  AUTHORS   Stevens,R.L.
  TITLE     Mast cell protease that cleaves fibrinogen
  JOURNAL   Patent: US 5968782-A 13 19-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .1137
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.98e-125       Length:         1137
Score:                  1402.00         Matches:        252
Percent Similarity:     95.24%          Conservative:   8
Best Local Similarity:  92.31%          Mismatches:     13
Query Match:            92.72%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982C-52 (1-275) x AR080460 (1-1137)

Qy          3 SerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaPro 22
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCA 62

Qy         23 ValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrp 42
                 ||||||||||||:::   |||||||||||||||||||||||||||||||||||||||
Db         63 GGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGG 122

Qy         43 ProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySer 62
              ||||||||||||||||||||||||         |||||||||||||||||||||||||||
Db        123 CCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCC 182

Qy         63 LeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAsp 82
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        183 CTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGAT 242

Qy         83 LeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeu 102
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTG 302

Qy        103 ProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIle 122
              ||||||||||||||||||||||||||||||||||||      |||   ||||||||||||
Db        303 CCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATC 362

Qy        123 AlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeu 142
              ||||||||||||||||||||||||||||||:::||||||   ||||||||||||   |||
Db        363 GCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTG 422

Qy        143 ProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAsp 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGAT 482

Qy        163 ValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIle 182
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        483 GTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATA 542

Qy        183 MetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspVal 202
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTC 602

Qy        203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGly 222
              ||||||:::|||||||||||||||||||||||||||::::::||||||||||||:::|||
Db        603 CGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGC 662

Qy        223 AspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValVal 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTC 722

Qy        243 SerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyr 262
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 AGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTAC 782

Qy        263 TyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||
Db        783 TACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 821

RESULT 15
HUMTRY1A
LOCUS       HUMTRY1A                1137 bp    mRNA    linear   PRI 03-AUG-1993
DEFINITION  Human tryptase-I mRNA, 3' end.
ACCESSION   M33491
VERSION     M33491.1  GI:339980
KEYWORDS    serine protease; tryptase-I.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1137)
  AUTHORS   Vanderslice,P., Ballinger,S.M., Tam,E.K., Goldstein,S.M.,
            Craik,C.S. and Caughey,G.H.
  TITLE     Human mast cell tryptase: multiple cDNAs and genes reveal a
            multigene serine protease family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 87 (10), 3811-3815 (1990)
  MEDLINE   90251647
   PUBMED   2187193
COMMENT     Original source text: Human adult skin, cDNA to mRNA.
            Draft entry and computer-readable sequence for [1] kindly submitted
            by P.Vanderslice, 02-APR-1990.
FEATURES             Location/Qualifiers
     source          1. .1137
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             <1. .824
                     /note="tryptase-I"
                     /codon_start=3
                     /protein_id="AAA36778.1"
                     /db_xref="GI:339981"
                     /translation="NLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQV
                     SLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPV
                     SRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGD
                     VDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSC
                     QGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
     polyA_signal    1119. .1124
ORIGIN

Alignment Scores:
Pred. No.:              2.98e-125       Length:         1137
Score:                  1402.00         Matches:        252
Percent Similarity:     95.24%          Conservative:   8
Best Local Similarity:  92.31%          Mismatches:     13
Query Match:            92.72%          Indels:         0
DB:                     9               Gaps:           0

US-09-598-982C-52 (1-275) x HUMTRY1A (1-1137)

Qy          3 SerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaPro 22
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCA 62
```

```
Qy        23 ValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrp 42
                ||||||||||||:::   |||||||||||||||||||||||||||||||||||||||
Db        63 GGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGG 122

Qy        43 ProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySer 62
             ||||||||||||||||||||||||         |||||||||||||||||||||||||||
Db       123 CCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCC 182

Qy        63 LeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAsp 82
             |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       183 CTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGAT 242

Qy        83 LeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeu 102
             ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       243 CTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTG 302

Qy       103 ProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIle 122
             ||||||||||||||||||||||||||||||||||||      |||   ||||||||||||
Db       303 CCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATC 362

Qy       123 AlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeu 142
             ||||||||||||||||||||||||||||||:::||||||   ||||||||||||   |||
Db       363 GCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTG 422

Qy       143 ProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAsp 162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       423 CCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGAT 482

Qy       163 ValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIle 182
             |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db       483 GTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATA 542

Qy       183 MetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspVal 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       543 ATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTC 602

Qy       203 ArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGly 222
             ||||||:::|||||||||||||||||||||||||||::::::||||||||||||:::|||
Db       603 CGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGC 662

Qy       223 AspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValVal 242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       663 GACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTC 722

Qy       243 SerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyr 262
             ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       723 AGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTAC 782

Qy       263 TyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
             |||||||||||||||||||||||||||||||||||||||
Db       783 TACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 821

Search completed: August 30, 2005, 18:42:00
Job time : 3564 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         August 30, 2005, 15:00:06 ; Search time 3116 Seconds
                                            (without alignments)
                                            3359.331 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters:      68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_29082005_080946_26105/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_5180_@runat_29082005_080946_26105 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  gb_est1:*
                2:  gb_est2:*
                3:  gb_htc:*
                4:  gb_est3:*
                5:  gb_est4:*
                6:  gb_est5:*
                7:  gb_est6:*
                8:  gb_gss1:*
                9:  gb_gss2:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                    Description
 ---------------------------------------------------------------------------
     1    1463   96.8   1030  5  BM919268              BM919268 AGENCOURT
     2  1204.5   79.7    904  5  BQ720404              BQ720404 AGENCOURT
     3    1129   74.7    619  6  CD671851              CD671851 fg07g11.y
     4    1119   74.0    631  7  CV030374              CV030374 9476 Full
     5  1116.5   73.8    673  6  CD616760              CD616760 56015240J
c    6  1098.5   72.7    682  6  CD616759              CD616759 56015240H
     7    1081   71.5    642  6  CD672689              CD672689 fg15a06.y
     8  1057.5   69.9    806  7  CF553058              CF553058 AGENCOURT
     9    1044   69.0    930  6  CB203717              CB203717 AGENCOURT
```

```
    10    1039   68.7    635  6  CD671891              CD671891 fg08b10.y
    11    1029   68.1   1000  4  BI413250              BI413250 602986390
    12    1006   66.5    589  5  BU071733              BU071733 im33a09.y
    13   989.5   65.4    811  4  BG697540              BG697540 602660760
    14     959   63.4    581  5  BU787772              BU787772 io34c08.y
    15     923   61.0    736  4  BG542089              BG542089 602571309
    16     905   59.9    678  4  BG483591              BG483591 602503309
c   17     900   59.5    647  6  CD631944              CD631944 56015309J
    18     897   59.3    722  5  BX091892              BX091892 BX091892
    19     891   58.9    512  2  BF724180              BF724180 bx01h02.y
    20     866   57.3    502  5  BQ082929              BQ082929 K-EST0144
    21     846   56.0    859  7  CF583025              CF583025 AGENCOURT
    22     830   54.9    671  6  CB840425              CB840425 M15E-0715
c   23     823   54.4    929  7  CF583026              CF583026 AGENCOURT
    24     818   54.1    639  1  AA131142              AA131142 zl31b01.r
    25   808.5   53.5    553  6  CD631946              CD631946 56015325J
    26     807   53.4    674  7  CK967791              CK967791 4083412 B
    27     796   52.6    917  5  BQ721078              BQ721078 AGENCOURT
c   28     767   50.7    686  5  BU073263              BU073263 im33a09.x
    29     742   49.1    558  4  BI790801              BI790801 id09c06.y
    30     714   47.2    719  5  BX922783              BX922783 BX922783
c   31     705   46.6    731  7  CK770530              CK770530 958715 MA
    32     697   46.1   1219  3  AK081986              AK081986 Mus muscu
    33     693   45.8    474  2  AW823937              AW823937 uf61e03.y
    34     691   45.7    391  1  AJ714198              AJ714198 AJ714198
    35   690.5   45.7    676  4  BG964077              BG964077 602828831
c   36     677   44.8    371  2  BF359130              BF359130 QV3-ET006
    37     675   44.6    620  1  AA049080              AA049080 mj50f09.r
c   38     667   44.1    688  5  BQ003361              BQ003361 UI-H-EI1-
c   39     662   43.8    693  5  BM991728              BM991728 UI-H-DF1-
    40     659   43.6    527  6  CD631943              CD631943 56015309H
c   41     657   43.5    699  6  CA438721              CA438721 UI-H-DT1-
    42     655   43.3    525  4  BI775716              BI775716 468400 MA
    43     654   43.3    545  4  BI681216              BI681216 460544 MA
    44     654   43.3    581  6  CA867461              CA867461 ir28f02.y
    45     651   43.1    433  7  CK826343              CK826343 io34c08.y


                                ALIGNMENTS


RESULT 1
BM919268
LOCUS       BM919268                1030 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6715837 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748499
            5', mRNA sequence.
ACCESSION   BM919268
VERSION     BM919268.1  GI:19369647
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1030)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12776  row: f  column: 20
            High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .1030
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5748499"
```

```
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_120"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                 7.64e-133      Length:        1030
Score:                     1463.00        Matches:       270
Percent Similarity:        98.92%         Conservative:  4
Best Local Similarity:     97.47%         Mismatches:    1
Query Match:               96.76%         Indels:        2
DB:                        5              Gaps:          0

US-09-598-982C-52 (1-275) x BM919268 (1-1030)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db         33 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGCGGCCCCT 92

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 152

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGTGGG 212

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 272

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 332

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 392

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 452

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 512

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 572

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 632

Qy        201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
              ||||||||||||||||||||||||||||||||||||||||||::::::||||||||||||
Db        633 GACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 692

Qy        221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 CAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTACAGGCGGGC 752

Qy        241 ValValSerTrp-AspGlu-GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgV 260
              |||||||||||| |||||| |||||||||||||||:::||||||||||||||||||||||
```

```
Db        753 GTGGTCAGCTGGGACGAGGGGCTGTGCCCAGCCCCACCGGCCTGGCATCTACACCCGTG 812

Qy        260 alThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro 275
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        813 TCACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCG 859

RESULT 2
BQ720404
LOCUS       BQ720404                 904 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8477973 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6196776 5', mRNA sequence.
ACCESSION   BQ720404
VERSION     BQ720404.1  GI:21859301
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13605  row: m  column: 01
            High quality sequence stop: 587.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6196776"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.9 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine); available through Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:                 1.5e-107       Length:        904
Score:                     1204.50        Matches:       225
Percent Similarity:        88.06%         Conservative:  11
Best Local Similarity:     83.96%         Mismatches:    21
Query Match:               79.66%         Indels:        11
DB:                        5              Gaps:          1

US-09-598-982C-52 (1-275) x BQ720404 (1-904)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 69

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         70 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 129
```

```
Qy          41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 189

Qy          61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
               |||||||||||||||||||||||||||||||||||||||||||||:::||||||      
Db         190 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCG------ 243

Qy          81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                                    |||||||||||||||||||||||||||||||||||||||
Db         244 ---------------------GTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 282

Qy         101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
               ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db         283 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 342

Qy         121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
               ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db         343 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 402

Qy         141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         403 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 462

Qy         161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         463 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 522

Qy         181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         523 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 582

Qy         201 AspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAspSerCys 220
               ||||||||||||:::|||||||||||||||||||||||||||::::::||||||||||||
Db         583 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGC 642

Qy         221 LysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGly 240
               :::||||||||||||||||||||||||         ||||||   ||||||   ||||||
Db         643 CAGGGCGACTCCGGAGGGCCCCTGGTGGGCAACGCGAATGGCCCCTGGCTGGCAGCGGGC 702

Qy         241 Val-ValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThr-ArgV 260
               ||| |||:::|||   ||||||   |||:::||||||||    |||   |||||| ||| 
Db         703 GTGGGTCACCTGGGGCGAGGGCTGGGCCCGGCCCAACCGGGCGGGCTTCTACACCCCGGG 762

Qy         260 alThrTyrTyrLeuAspTrp 266
                       |||   :::|||
Db         763 CCCCGTCCTACTGGGAATGG 782

RESULT 3
CD671851
LOCUS       CD671851                 619 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg07g11.y1 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg07g11 5', mRNA sequence.
ACCESSION   CD671851
VERSION     CD671851.1  GI:32173582
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
   MEDLINE  22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 07  row: g  column: 11
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..619
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg07g11"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              2.27e-100       Length:          619
Score:                  1129.00         Matches:         202
Percent Similarity:     100.00%         Conservative:    3
Best Local Similarity:  98.54%          Mismatches:      0
Query Match:            74.67%          Indels:          0
DB:                     6               Gaps:            0

US-09-598-982C-52 (1-275) x CD671851 (1-619)

Qy          34 GlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgValArgAspArg 53
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 GGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGA 62

Qy          54 TyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAla 73
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCG 122

Qy          74 HisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgValGlnLeuArgGluGln 93
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 CACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAG 182

Qy          94 HisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHisProGlnPhe 113
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 CACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTC 242

Qy         114 TyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGluGluProValAsnIle 133
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 TACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATC 302

Qy         134 SerSerArgValHisThrValMetLeuProProAlaSerGluThrPheProProGlyMet 153
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 TCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATG 362

Qy         154 ProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluProLeuProProProPhe 173
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 CCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTT 422

Qy         174 ProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAlaLysTyrHis 193
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         423 CCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCAC 482
```

```
Qy          194 LeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAspMetLeuCysAlaGly 213
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          483 CTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGG 542

Qy          214 AsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValCysLysValAsn 233
                |||::::::||||||||||||:::||||||||||||||||||||||||||||||||||||
Db          543 AACACCCGGAGGGACTCATGCCAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAAT 602

Qy          234 GlyThrTrpLeuGln 238
                |||||||||||||||
Db          603 GGCACCTGGCTACAG 617

RESULT 4
CV030374
LOCUS       CV030374                 631 bp    mRNA    linear   EST 20-AUG-2004
DEFINITION  9476 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC028059, mRNA sequence.
ACCESSION   CV030374
VERSION     CV030374.1  GI:51488624
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGCTGAGCCTGCTGCTG
            BACKWARD: CACGGCTTTTTGGGGACATA
            Insert Length: 631    Std Error: 27.00
            Plate: 11081   row: 06   column: C
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 98
            High quality sequence stop: 630
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="mixed"
                     /clone_lib="Full Length cDNA from the Mammalian Gene
                     Collection"
                     /note="Vector: mixed; The ORFs were PCR amplified from the
                     MGC (Mammalian Gene Collection) as of April 2004 and
                     cloned by recombinational Gateway cloning into pDONR223
                     Donor vector. Reference : MGC (Mammalian Gene Collection)
                     Program Team, Generation and Initial Analysis of more than
                     15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                     2002, 99(26), 16899-16903"
ORIGIN

Alignment Scores:
Pred. No.:             2.24e-99       Length:          631
Score:                 1119.00        Matches:         208
Percent Similarity:    99.05%         Conservative:    0
```

```
Best Local Similarity:   99.05%         Mismatches:     2
Query Match:             74.01%         Indels:         0
DB:                      7              Gaps:           0

US-09-598-982C-52 (1-275) x CV030374 (1-631)

Qy            1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
                ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db            1 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGCGGCCCCT 60

Qy           21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 120

Qy           41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGTGGG 180

Qy           61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 240

Qy           81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 300

Qy          101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 360

Qy          121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 420

Qy          141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 480

Qy          161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTC 540

Qy          181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
                ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db          541 CCCATAATGGAAAACCACATTTGTGACGCANAATACCACCTTGGCGCCTACACGGGAGAC 600

Qy          201 AspValArgIleIleArgAspAspMetLeu 210
                ||||||||||||||||||||||||||||||
Db          601 GACGTCCGCATCATCCGTGACGACATGCTG 630

RESULT 5
CD616760
LOCUS       CD616760                 673 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56015240J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD616760
VERSION     CD616760.1  GI:40265024
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 673)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .673
```

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:              4.31e-99       Length:         673
Score:                  1116.50        Matches:        216
Percent Similarity:     96.00%         Conservative:   0
Best Local Similarity:  96.00%         Mismatches:     8
Query Match:            73.84%         Indels:         4
DB:                     6              Gaps:           0

US-09-598-982C-52 (1-275) x CD616760 (1-673)

Qy         12 LeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGlnAlaGlyIle 31
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATC 61

Qy         32 ValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgValArg 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGC 121

Qy         52 AspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeuThr 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACC 181

Qy         72 AlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgValGlnLeuArg 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTGCGG 241

Qy         92 GluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHisPro 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCACCCA 301

Qy        112 GlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGluGluProVal 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCCGTG 361

Qy        132 AsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThrPheProPro 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCG 421

Qy        152 GlyMetProCysTrpValThrGlyTrp-GlyAspValAspAsnAspGluProLeuProPr 171
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        422 GGGATGCCGTGCTGGGTCACTGGCTGGGGGCGATGTGGACAATGATGAGCCCCTCCCACC 481

Qy        171 oProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAlaLy 191
              |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GCAATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAA 541

Qy        191 sTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAspMetLeuCy 211
              ||||||||||      ||||||||||||||||||||||||||||||||||||||||||||
Db        542 ATACCACCTTTGGCGCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTGTG 601

Qy        211 sAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValCysLy 231
              ||||||||||   |||||||||   ||||||||||||||||||||||||      |||
Db        602 TGCCGGGAACAG-CAGAGGGACTC-TGCAAGGGCGACTCTGGAGGGCCCTGGTCTTGCAG 659

Qy        231 sValAsnGlyThr 235
               ||||||||||||
Db        660 -GTGAATGGCACC 671

RESULT 6
CD616759/c
LOCUS       CD616759                 682 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56015240H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD616759
VERSION     CD616759.1  GI:40265023
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1..682
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:              2.56e-97       Length:         682
Score:                  1098.50        Matches:        221
Percent Similarity:     97.37%         Conservative:   1
Best Local Similarity:  96.93%         Mismatches:     5
Query Match:            72.65%         Indels:         6
DB:                     6              Gaps:           0

US-09-598-982C-52 (1-275) x CD616759 (1-682)

Qy          9 LeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeuGlnGln 28
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        678 CTGCCCGTCCTGGCGAC-CGCGCCTACGCGGCGCCTGCCCCAGTCCAGGCCCTGCAGCAA 620

Qy         29 AlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeu 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 GCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTG 560

Qy         49 ArgValArgAspArgTyrTrpMet-HisPheCysGlyGlySerLeuIleHisProGlnTr 68
              |||||||||||||||||||||::: ||||||||||||||||||||||||||||||   ||
Db        559 AGAGTCCGCGACCGATACTGGATAGCACTTCTGCGGGGGCTCCCTCATCCACCCCAG-TG 501

Qy         68 pValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgVa 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGT 441

Qy         88 lGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIl 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCAT 381

Qy        108 eValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGl 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CGTGCACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGA 321

Qy        128 uGluProValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluTh 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGAC 261

Qy        148 rPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluPr 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCC 201

Qy        168 oLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCy 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTG 141

Qy        188 sAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAs 208
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 TGACGC-AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGA 82

Qy        208 pMetLeuCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLe 228
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 CATGCTGTGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCTG 22

Qy         228 uValCysLysValAsnGlyThr 235
                ||||||   ||||||||||||
Db          21 -GTCTGCAG-GTGAATGGCACC 2

RESULT 7
CD672689
LOCUS       CD672689                 642 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg15a06.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg15a06 5', mRNA sequence.
ACCESSION   CD672689
VERSION     CD672689.1  GI:32174420
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
  MEDLINE   22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 15  row: a  column: 06
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg15a06"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              1.23e-95      Length:         642
Score:                  1081.00       Matches:        198
Percent Similarity:     95.75%        Conservative:   5
Best Local Similarity:  93.40%        Mismatches:     9
Query Match:            71.49%        Indels:         0
DB:                     6             Gaps:           0

US-09-598-982C-52 (1-275) x CD672689 (1-642)

Qy           1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
               ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db           6 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 65

Qy          21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
               ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db          66 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 125

Qy          41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         126 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGCGGG 185

Qy          61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
               |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db         186 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 245

Qy          81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
               ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db         246 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 305

Qy         101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
               ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db         306 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 365

Qy         121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
               ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db         366 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 425

Qy         141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 GCCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 485

Qy         161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         486 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 545

Qy         181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         546 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 605

Qy         201 AspValArgIleIleArgAspAspMetLeuCysAla 212
               ||||||||||||:::|||||||||||||||||||||
Db         606 GACGTCCGCATCGTCCGTGACGACATGCTGTGTGCC 641

RESULT 8
CF553058
LOCUS       CF553058                 806 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15594982 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529468 5', mRNA sequence.
ACCESSION   CF553058
VERSION     CF553058.1  GI:34889892
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 806)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM617  row: b  column: 05
```

```
            High quality sequence start: 18
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1. .806
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529468"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

Alignment Scores:
Pred. No.:              3.42e-93      Length:        806
Score:                  1057.50       Matches:       202
Percent Similarity:     95.75%        Conservative:  1
Best Local Similarity:  95.28%        Mismatches:    6
Query Match:            69.94%        Indels:        3
DB:                     7             Gaps:          1

US-09-598-982C-52 (1-275) x CF553058 (1-806)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        158 ATGCTGAGCCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCCGGCCTACGCGGCCCCT 217

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 GCCCCAGTCCAGGCCCTGCAGCAAGCGGGTATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 277

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGATACTGGATGCACTTCTGTGGG 337

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCGCACTGCCTGGGACCGGACGTC 397

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 AAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 457

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACATCATCCAGACTGGAGCG 517

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 GATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCGCGTCCACACGGTC 577

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 ATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 637

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 GGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAAGCAGGTGAAAGTC 697

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGly--AlaTyrThrGlyA 200
              |||||||||||||||||||||||||||||||||||||||||||||     ||||||||||
Db        698 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGGCGCCCTACACGGGAG 757

Qy        200 spAsp---ValArgIleIleArgAspAspMet 209
              ||:::               ||||||||||||
Db        758 ACGAACGTCCGGCATCATCCCGTGACGACATG 789

RESULT 9
CB203717
```

```
LOCUS       CB203717               930 bp    mRNA    linear   EST 05-FEB-2003
DEFINITION  AGENCOURT_11288628 NIH_MGC_135 Mus musculus cDNA clone
            IMAGE:30142408 5', mRNA sequence.
ACCESSION   CB203717
VERSION     CB203717.1  GI:28240351
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM0051  row: j  column: 17
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30142408"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_135"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRV; Site_2: NotI;
                     Normalized full-length enriched library from pooled mouse
                     embryonic limb, maxilla and mandible, day 12.5, 13.5,
                     14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
                     Cloned directionally, priming method: Oligo-dT. cDNA
                     enrichment: >1k bp, Average insert size 1.6k bp.
                     Normalization (Cot value): 7.5 kb. Priming sequence:
                     5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)3' Tissue contributed
                     by, David Rowe. Library constructed by ResGen, Invitrogen
                     Corp."
ORIGIN

Alignment Scores:
Pred. No.:              8.82e-92      Length:        930
Score:                  1044.00       Matches:       189
Percent Similarity:     82.76%        Conservative:  27
Best Local Similarity:  72.41%        Mismatches:    39
Query Match:            69.05%        Indels:        6
DB:                     6             Gaps:          1

US-09-598-982C-52 (1-275) x CB203717 (1-930)

Qy         17 TyrAlaAlaProAlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGlu 36
              |||:::||||||   |||      |||   |||:::   |||||||||||||||   |||
Db          6 TACTCAGCCCCTCGCCCA------GCCAATCAGCGAGTGGGCATCGTGGAGGACATGAG 59

Qy         37 AlaProArgSerLysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMet 56
              |||      ||||||||||||||||||||||||||||||   :::      ||||||:::
Db         60 GCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTTAAATTAAACTACTGGATA 119

Qy         57 HisPheCysGlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeu 76
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        120 CATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTCACTGCGGCACACTGTGTG 179

Qy         77 GlyProAspValLysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyr 96
              ||||||   :::|||               |||||||||||||||||||||:::||||||
Db        180 GGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTTCGTGAGCAGTATCTATAC 239

Qy         97 TyrGlnAspGlnLeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIle 116
              |||   ||||||||||||   ::::::||||||:::|||||||||   :::|||
```

```
Db        240 TATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCACCCCCACTATTACACGGCC 299

Qy        117 GlnThrGlyAlaAspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArg 136
              :::   |||||||||:::||||||||||||||||||   |||||||||:::|||:::
Db        300 GAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCTGTGAATGTCTCCACCCAT 359

Qy        137 ValHisThrValMetLeuProProAlaSerGluThrPheProProGlyMetProCysTrp 156
              :::|||   :::   |||||||||||||||||||||||||||||||||      ||||||
Db        360 ATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCCCCTGGGACATCGTGCTGG 419

Qy        157 ValThrGlyTrpGlyAspValAspAsnAspGluProLeuProProProPheProLeuLys 176
              ||||||||||||||||||:::|||||||||||||||||||||||||||:::|||||||||
Db        420 GTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCACCTCCTTATCCTCTGAAG 479

Qy        177 GlnValLysValProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAla 196
              ||||||||||||||||||:::||||||   :::||||||   |||||||||   |||
Db        480 CAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGGAAGTACCACACTGGCCTC 539

Qy        197 TyrThrGlyAspAspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGln 216
              |||||||||||||||      |||:::   |||   ||||||||||||||||||::::::
Db        540 TACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTGTGTGCTGGAAATACCAGG 599

Qy        217 ArgAspSerCysLysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrp 236
              ||||||||||||:::|||||||||||||||||||||||||||||||||   |||||||||
Db        600 AGAGACTCCTGCCAGGGCGATTCAGGGGGCCACTGGTCTGCAAAGTGAAGGGTACCTGG 659

Qy        237 LeuGlnAlaGlyValValSerTrp-AspGluGlyCysAlaGlnProAsnArgProGlyIl 256
              ||||||||||||||||||||||||    |||||||||||||||||||||:::||||||||
Db        660 CTGCAGGCAGGAGTGGTCAGCTGGGNGTGAGGGCTGCGCACAGCCCAACAAGCCTGGCAT 719

Qy        256 eTyrThrArg-ValThr-TyrTyrLeuAspTrpIleHis-HisTyrValProLys 273
              |||||||||| ||||||    ||||||||||||||||||    |||||||||:::
Db        720 CTACACCCGGGGTGACATANCTACTTAGACTGGATCCACCNGCTATGTCCCTGAA 774

RESULT 10
CD671891
LOCUS       CD671891                 635 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg08b10.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg08b10 5', mRNA sequence.
ACCESSION   CD671891
VERSION     CD671891.1  GI:32173622
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
   MEDLINE  22103462
    PUBMED  12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 08  row: b  column: 10
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .635
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg08b10"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              1.61e-91        Length:         635
Score:                  1039.00         Matches:        191
Percent Similarity:     93.78%          Conservative:   5
Best Local Similarity:  91.39%          Mismatches:     13
Query Match:            68.72%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982C-52 (1-275) x CD671891 (1-635)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCT 66

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         67 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGC 126

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        127 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCATGGCCCATACTGGATGCACTTCTGCGGG 186

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              ||||||||||||||||||||||||||||||   ||||||||||||:::||||||||||||
Db        187 GGCTCCCTCATCCACCCCCAGTGGGTGCTGCTCGCAGCGCACTGCGTGGGACCGGACGTC 246

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        247 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 306

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        307 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCCCAGTTCTACACCGCCCAGATCGGAGCG 366

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              ||||||||||||||||||||||||||||||||||||:::||||||   ||||||||||||
Db        367 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTC 426

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 486

Qy        161 GlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnValLysVal 180
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        487 GGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTC 546

Qy        181 ProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 CCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGAC 606

Qy        201 AspValArgIleIleArgAspAspMet 209
              ||||||||||||:::||||||||||||
Db        607 GACGTCCGCATCGTCCGTGACGACATG 633

RESULT 11
BI413250
```

```
LOCUS       BI413250                1000 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  602986390F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5142478 5',
            mRNA sequence.
ACCESSION   BI413250
VERSION     BI413250.1  GI:15174173
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1000)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by:Incyte Genomics, Inc.
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11351  row: c  column: 23
            High quality sequence start: 32
            High quality sequence stop: 849.
FEATURES             Location/Qualifiers
     source          1..1000
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="CZECH II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5142478"
                     /tissue_type="pooled lung tumors"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NCI_CGAP_Lu33"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was prepared from mRNA obtained from pooled
                     lung tumors with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCTGTTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites of  the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Alignment Scores:
Pred. No.:              2.9e-90       Length:        1000
Score:                  1029.00       Matches:       189
Percent Similarity:     81.27%        Conservative:  28
Best Local Similarity:  70.79%        Mismatches:    43
Query Match:            68.06%        Indels:        7
DB:                     4             Gaps:          1

US-09-598-982C-52 (1-275) x BI413250 (1-1000)

Qy          11 ValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGlnAlaLeu-GlnGlnAlaGl 30
               :::      |||      |||:::||||||   |||      |||    |||:::   ||
Db          20 ATTTANCTTAGTCTGGTGTACTCAGCCCCTCGCCCA------GCCAAATCAGCGAGTGGG 73

Qy          30 yIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVa 50
               |||||||||||||   ||||||      ||||||||||||||||||||||||||||||
Db          74 CATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATT 133

Qy          50 lArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLe 70
                :::      ||||||:::|||||||||||||||||||||||||||||||||||||||||
Db         134 TAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCT 193

Qy          70 uThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgValGlnLe 90
               ||||||||||||||||:::||||||   :::|||               |||||||||||
Db         194 CACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCT 253

Qy          90 uArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHi 110
               ||||||||||:::|||||||||   ||||||||||||   ::::::||||||:::|||||
Db         254 TCGTGAGCAGTATCTATACTATGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCA 313

Qy         110 sProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuLeuGluLeuGluGluPr 130
               ||||   :::|||      :::   |||||||||:::||||||||||||||||||   ||
Db         314 CCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCC 373

Qy         130 oValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGluThrPhePr 150
               |||||||:::|||:::   :::|||   :::   ||||||||||||||||||||||||||
Db         374 TGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCC 433

Qy         150 oProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluProLeuPr 170
               |||||||      ||||||||||||||||||||||||:::||||||||||||||||||||
Db         434 CCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCC 493

Qy         170 oProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAl 190
               |||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db         494 ACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCG 553

Qy         190 aLys-TyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleIleArgAspAspMetL 210
                |||    |||   |||   |||||||||||||||      |||:::   |||   ||||
Db         554 GAAGTAACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGC 613

Qy         210 euCysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValC 230
               ||||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||
Db         614 TGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCT 673

Qy         230 ysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSer-TrpAspGluGlyCysAla 249
               ||||||||   ||||||||||||||||||||||||||||::: |||   ||||||||||||
Db         674 GCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCACGCTGGGGTGAGGGCTGCGCA 733

Qy         250 Gln-ProAsnArgProGlyIleTyrThrArgValThr-TyrTyrLeuAspTrpIleHisH 269
               ||| ||||||:::|||||||||||||||||||||||| |||||||||||||||||||||
Db         734 CAGCCCAAACAAGCCTGGCATCTACACCCGGGTGACCATACTACTTAGACTGGATCCACG 793

Qy         269 isTyrValPro 272
                 |||||||||
Db         794 CTTATGTCCCT 804

RESULT 12
BU071733
LOCUS       BU071733                 589 bp    mRNA    linear   EST 27-AUG-2002
DEFINITION  im33a09.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6036761
            5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   BU071733
VERSION     BU071733.1  GI:22512922
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: im33a09.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
```

```
                  Library was constructed by Dr. J. Ferrer In vivo mass-excised to
                  pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
                  University Genome Sequencing Center For information on obtaining a
                  clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
                  Seq primer: -40RP from Gibco
                  High quality sequence stop: 433.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6036761"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:              2.51e-88       Length:         589
Score:                  1006.00        Matches:        180
Percent Similarity:     95.85%         Conservative:   5
Best Local Similarity:  93.26%         Mismatches:     8
Query Match:            66.53%         Indels:         0
DB:                     5              Gaps:           0

US-09-598-982C-52 (1-275) x BU071733 (1-589)

Qy         79 AspValLysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGln 98
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db         10 GACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAG 69

Qy         99 AspGlnLeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThr 118
              ||||||||||||||||||||||||||||||||||||||||||||||||      |||
Db         70 GACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATC 129

Qy        119 GlyAlaAspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHis 138
              ||||||||||||||||||||||||||||||||||||||||||:::||||||   ||||||
Db        130 GGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCAC 189

Qy        139 ThrValMetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThr 158
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACT 249

Qy        159 GlyTrpGlyAspValAspAsnAspGluProLeuProProProPheProLeuLysGlnVal 178
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        250 GGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTG 309

Qy        179 LysValProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThr 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 AAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACG 369

Qy        199 GlyAspAspValArgIleIleArgAspAspMetLeuCysAlaGlyAsnSerGlnArgAsp 218
              ||||||||||||||||||:::|||||||||||||||||||||||||||::::::||||||
Db        370 GGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGAC 429

Qy        219 SerCysLysGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGln 238
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 TCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAG 489

Qy        239 AlaGlyValValSerTrpAspGluGlyCysAlaGlnProAsnArgProGlyIleTyrThr 258
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        490 GCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACC 549

Qy        259 ArgValThrTyrTyrLeuAspTrpIleHisHisTyrVal 271
              |||||||||||||||||||||||||||||||||||||||
Db        550 CGTGTCACCTACTACTTGGACTGGATCCACCACTATGTC 588

RESULT 13
BG697540
LOCUS       BG697540                 811 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602660760F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4803879 5',
            mRNA sequence.
ACCESSION   BG697540
VERSION     BG697540.1  GI:13963867
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10700  row: g  column: 16
            High quality sequence stop: 781.
FEATURES             Location/Qualifiers
     source          1. .811
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4803879"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.63e-86       Length:         811
Score:                  989.50         Matches:        200
Percent Similarity:     77.41%         Conservative:   9
Best Local Similarity:  74.07%         Mismatches:     17
Query Match:            65.44%         Indels:         45
DB:                     4              Gaps:           3

US-09-598-982C-52 (1-275) x BG697540 (1-811)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaPro 20
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||   |||
Db         11 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGC-CCT 69

Qy         21 AlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSer 40
              ||||||   ||||||||||||:::   |||||||||||||||||||||||||||||||||
Db         70 GCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAGC 129

Qy         41 LysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGly 60
              ||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db        130 AAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGG 189

Qy         61 GlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        190 GGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTC 249

Qy         81 LysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGln 100
```

```
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        250 AAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAG 309

Qy        101 LeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAla 120
              ||||||||||||||||||||||||||||||||||||||||||      |||   ||||||
Db        310 CTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCG 369

Qy        121 AspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVal 140
              |||||||||||||||||||||||||||||||||   :::||||||   ||||||||||||
Db        370 GACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTC 429

Qy        141 MetLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrp 160
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 ACCCTGCCCCCTGCCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 489

Qy        161 GlyAspValAspAsnAsp------------------------------------------ 166
              ||||||||||||||||||
Db        490 GGCGATGTGGACAATGATGGTGGGTCTGGGGACAGTGGAGGTGGGGCCAGGGTCTTAGCC 549

Qy        166 ------------------------------------------------------------ 166

Db        550 ACAGCCCAGCCCCTGGGCTCCCTCTGGGCTCCAGGTGGGGGTTGCCCGGCCCCTCCTGAG 609

Qy        167 ---GluProLeuPro---------------Pro-ProPheProLeuLysGlnValLysVa 180
                 :::|||||||||               |||    |||||||||||||||||||||||
Db        610 GCTGACCCTCTTCCCCACTGCAGAGCGCTCCCACGCATTTCCTCTGAAGCAGGTGAAGGT 669

Qy        180 lProIleMetGluAsnHisIleCysAspAlaLysTyrHis-LeuGlyAlaTyrThrGlyA 200
              ||||||||||||||||||||||||||||   |||   ||| |||||||||||||||||||
Db        670 CCCCATAATGGAAAACCACATTTGTGACGGCAAGATACACCTTGGGCGCCTACACGGGAG 729

Qy        200 spAspValArgIleIleArgAspAsp-MetLeuCysAlaGlyAsnSerGlnArgAspSer 219
              ||||||||||||||:::||||||::: ||||||||||||||||||   :::   ||||||
Db        730 ACGACGTCCGCATCGTCCGTGACGAACATGCTGTGTGCGGGAAACCCGGAG---GACTCA 786

Qy        220 CysLysGlyAspSerGlyGlyPro 227
              |||:::||||||||||||||||||
Db        787 TGCCAGGGCGACTCCGGAGGGCCC 810

RESULT 14
BU787772
LOCUS       BU787772                581 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  io34c08.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128247
            5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   BU787772
VERSION     BU787772.1  GI:23835713
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: io34c08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6128247"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:              1.01e-83        Length:         581
Score:                  959.00          Matches:        176
Percent Similarity:     93.72%          Conservative:   3
Best Local Similarity:  92.15%          Mismatches:     12
Query Match:            63.43%          Indels:         0
DB:                     5               Gaps:           0

US-09-598-982C-52 (1-275) x BU787772 (1-581)

Qy          5 LeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaAlaProAlaProValGln 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db          8 CTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGGCCCCTGCCCCAGGCCAG 67

Qy         25 AlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrp 44
              |||||||||:::   |||||||||||||||||||||||||||||||||||||||||||||
Db         68 GCCCTGCAGCGAGTGGGCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 127

Qy         45 GlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIle 64
              ||||||||||||||||||         |||||||||||||||||||||||||||||||||
Db        128 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 187

Qy         65 HisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAla 84
              |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        188 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 247

Qy         85 ThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProVal 104
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 307

Qy        105 SerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeu 124
              ||||||||||||||||||||||||||||||      |||   ||||||||||||||||||
Db        308 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 367

Qy        125 LeuGluLeuGluGluProValAsnIleSerSerArgValHisThrValMetLeuProPro 144
              ||||||||||||||||||||||||:::||||||   ||||||||||||   |||||||||
Db        368 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 427

Qy        145 AlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAsp 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 487

Qy        165 AsnAspGluProLeuProProProPheProLeuLysGlnValLysValProIleMetGlu 184
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 547

Qy        185 AsnHisIleCysAspAlaLysTyrHisLeuGly 195
              |||||||||||||||||||||||||||||||||
```

```
Db        548 AACCACATTTGTGACGCAAAATACCACCTTGGC 580

RESULT 15
BG542089
LOCUS       BG542089                 736 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602571309F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4695458 5',
            mRNA sequence.
ACCESSION   BG542089
VERSION     BG542089.1  GI:13534322
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1521  row: b  column: 03
            High quality sequence stop: 725.
FEATURES             Location/Qualifiers
     source          1. .736
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4695458"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              4.79e-80       Length:         736
Score:                  923.00         Matches:        199
Percent Similarity:     85.95%         Conservative:   9
Best Local Similarity:  82.23%         Mismatches:     24
Query Match:            61.04%         Indels:         12
DB:                     4              Gaps:           1

US-09-598-982C-52 (1-275) x BG542089 (1-736)

Qy          1 MetLeuSerLeuLeuLeuLeuAlaLeuProValLeuAlaSerArg-AlaTyrAlaAlaPr 20
              ||||||:::|||||||||||||||||||||||||||||||||    ||||||||||||||
Db         24 ATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCTGCGCCTACGCGGCCCC 83

Qy         20 oAlaProValGlnAlaLeuGlnGlnAlaGlyIleValGlyGlyGlnGluAlaProArgSe 40
              |||||||   ||||||||||||:::   ||||||||||||||||||||||||||||||||
Db         84 TGCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCAGGAG 143

Qy         40 rLysTrpProTrpGlnValSerLeuArgValArgAspArgTyrTrpMetHisPheCysGl 60
              |||||||||||||||||||||||||||||||         ||||||||||||||||||||
Db        144 CAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGG 203

Qy         60 yGlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysLeuGlyProAspVa 80
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db        204 GGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGT 263

Qy         80 lLysAspLeuAlaThrLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGl 100
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        264 CAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCA 323

Qy        100 nLeuLeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrGlyAl 120
              |||||||||||||||||||||||||||||||||||||||||||      |||   |||||
Db        324 GCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGC 383

Qy        120 aAspIleAlaLeuLeuGluLeuGluGluProValAsnIleSerSerArgValHisThrVa 140
              ||||||||||||||||||||||||||||||||||   :::||||||   |||||||||||
Db        384 GGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGT 443

Qy        140 lMetLeuProProAlaSerGluThrPheProPro-GlyMetProCysTrpValThrGlyT 160
              |   |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        444 CACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCT 503

Qy        160 rpGlyAspValAsp-AsnAspGluProLeuProProProPheProLeuLysGlnVal-Ly 179
              |||||||||||||| |||||||||      |||      |||   |||||||||||| ||
Db        504 GGGGCGATGTGGACCAATGATGAGCGCCT-CCACGGCATTTCCTTCTGAAGCAGGTTGAA 562

Qy        179 sValProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAla-TyrThrG 199
              |||||||||||||||||||||||||||||||||||||   ||||||||||||    ||||
Db        563 GGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATAACACCTTGGGGCCTAACACGG 622

Qy        199 lyAspAsp-ValArgIle------IleArgAspAspMetLeuCysAlaGlyAsnSerGln 216
              |||||::: |||||||||      :::   |||         ||||||||||||::::::
Db        623 GAGACGAAGGTCCGCATTCGTCCGGTGAACGACAT-TGCTGTTGTGCCGGGAACACCCGG 681

Qy        217 ArgAspSerCysLysGly-AspSer-GlyGlyProLeuValCysLysVal 232
              ||||||||||||:::||| |||||| ||||||||||||||||||||||||
Db        682 AGGGACTCATGCCAGGGCGGACTCCGGGAGGGCCCCTGGTGTGCAAGGTG 731

Search completed: August 30, 2005, 17:42:35
Job time : 3128 secs
```

```
                      GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         August 27, 2005, 23:29:16 ; Search time 114 Seconds
                                           (without alignments)
                                           947.848 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters:      1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                16:  /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
                17:  /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
                18:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                19:  /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
                20:  /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
                21:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                22:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result        Query
  No.  Score  Match Length DB  ID                     Description
-------------------------------------------------------------------------
    1   1512  100.0    275  14  US-10-117-323-36        Sequence 36, Appl
    2   1512  100.0    275  15  US-10-352-684A-46       Sequence 46, Appl
    3   1414   93.5    275  15  US-10-287-226-90        Sequence 90, Appl
    4   1411   93.3    275  15  US-10-287-226-92        Sequence 92, Appl
    5   1411   93.3    275  16  US-10-723-860-2690      Sequence 2690, Ap
    6   1411   93.3    275  18  US-10-756-149-5451      Sequence 5451, Ap
    7   1405   92.9    275  10  US-09-813-432-44        Sequence 44, Appl
    8   1405   92.9    275  15  US-10-352-684A-48       Sequence 48, Appl
    9   1405   92.9    275  15  US-10-174-364-44        Sequence 44, Appl
   10   1405   92.9    275  15  US-10-246-583-44        Sequence 44, Appl
   11   1405   92.9    275  15  US-10-287-226-94        Sequence 94, Appl
```

```
   12   1405   92.9    275  16  US-10-689-832-44        Sequence 44, Appl
   13   1401   92.7    691  15  US-10-275-505-13        Sequence 13, Appl
   14 1386.5   91.7    272  14  US-10-117-323-40        Sequence 40, Appl
   15   1373   90.8    245  11  US-09-789-210-69        Sequence 69, Appl
   16   1360   89.9    264  14  US-10-131-409-104       Sequence 104, App
   17   1360   89.9    264  15  US-10-139-854-104       Sequence 104, App
   18   1360   89.9    264  15  US-10-150-813-104       Sequence 104, App
   19   1360   89.9    264  15  US-10-150-811-104       Sequence 104, App
   20 1356.5   89.7    279  15  US-10-287-226-88        Sequence 88, Appl
   21   1349   89.2    266  14  US-10-131-409-84        Sequence 84, Appl
   22   1349   89.2    266  15  US-10-139-854-84        Sequence 84, Appl
   23   1349   89.2    266  15  US-10-150-813-84        Sequence 84, Appl
   24   1349   89.2    266  15  US-10-150-811-84        Sequence 84, Appl
   25   1324   87.6    256  14  US-10-131-409-92        Sequence 92, Appl
   26   1324   87.6    256  15  US-10-139-854-92        Sequence 92, Appl
   27   1324   87.6    256  15  US-10-150-813-92        Sequence 92, Appl
   28   1324   87.6    256  15  US-10-150-811-92        Sequence 92, Appl
   29   1313   86.8    254  14  US-10-131-409-74        Sequence 74, Appl
   30   1313   86.8    254  15  US-10-139-854-74        Sequence 74, Appl
   31   1313   86.8    254  15  US-10-150-813-74        Sequence 74, Appl
   32   1313   86.8    254  15  US-10-150-811-74        Sequence 74, Appl
   33   1186   78.4    231  10  US-09-898-837A-52       Sequence 52, Appl
   34   1134   75.0    310  15  US-10-275-505-8         Sequence 8, Appli
   35   1133   74.9    276  14  US-10-117-323-35        Sequence 35, Appl
   36   1118   73.9    273  15  US-10-028-248A-63       Sequence 63, Appl
   37   1118   73.9    273  15  US-10-107-782-63        Sequence 63, Appl
   38   1113   73.6    276  14  US-10-311-955-3         Sequence 3, Appli
   39   1108   73.3    242  17  US-10-480-988-27        Sequence 27, Appl
   40   1103   72.9    273  15  US-10-028-248A-64       Sequence 64, Appl
   41   1103   72.9    273  15  US-10-107-782-64        Sequence 64, Appl
   42   1094   72.4    252  10  US-09-813-432-42        Sequence 42, Appl
   43   1094   72.4    252  15  US-10-174-364-42        Sequence 42, Appl
   44   1094   72.4    252  15  US-10-246-583-42        Sequence 42, Appl
   45   1094   72.4    252  16  US-10-689-832-42        Sequence 42, Appl


                              ALIGNMENTS


RESULT 1
US-10-117-323-36
; Sequence 36, Application US/10117323
; Publication No. US20030054993A1
; GENERAL INFORMATION:
;  APPLICANT: Rancourt, Derrick E.
;  APPLICANT:  Rancourt, Susan L.
;  APPLICANT:  O'Sullivan, Colleen M.
;  TITLE OF INVENTION: Implantation Serine Proteinases
;  FILE REFERENCE: 033337-005
;  CURRENT APPLICATION NUMBER: US/10/117,323
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US 60/281,724
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/294,736
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: US 60/350,962
;  PRIOR FILING DATE: 2002-01-25
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-117-323-36

  Query Match             100.0%;  Score 1512;  DB 14;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-130;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
```

```
Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 2
US-10-352-684A-46
; Sequence 46, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-352-684A-46

  Query Match            100.0%;  Score 1512;  DB 15;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-130;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 3
US-10-287-226-90
; Sequence 90, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
```

```
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 90
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-287-226-90

  Query Match             93.5%;  Score 1414;  DB 15;  Length 275;
  Best Local Similarity   92.7%;  Pred. No. 4.3e-121;
  Matches  255;  Conservative    8;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              |||||||||||||| || |||| ||||| ||||||||||||||||||||||  |||||||
Db          1 MLSLLLLALPVLASPAYVAPAPGQALQQTGIVGGQEAPRSKWPWQVSLRVRGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||: ||:|||| |||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCVEPDIKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||| :||| ||||||||||||||||||||||||:| ||||||||||||
Db        121 DIALLELEEPVNISSHIHTVTLPPASETFPPGMPCWVTGWGDVDNNERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 4
US-10-287-226-92
; Sequence 92, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
```

```
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 92
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-287-226-92

  Query Match             93.3%;  Score 1411;  DB 15;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 8.1e-121;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 5
US-10-723-860-2690
; Sequence 2690, Application US/10723860
```

```
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2690
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-860-2690

  Query Match             93.3%;  Score 1411;  DB 16;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 8.1e-121;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275

RESULT 6
US-10-756-149-5451
; Sequence 5451, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5451
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-756-149-5451

  Query Match             93.3%;  Score 1411;  DB 18;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 8.1e-121;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275

RESULT 7
US-09-813-432-44
; Sequence 44, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier Jr., Raymond J
;  APPLICANT:  Majmuder, Kamud
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Smithson, Glenda
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Vernet, Corine A. M.
;  TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
;  FILE REFERENCE: 15966-729
;  CURRENT APPLICATION NUMBER: US/09/813,432
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 60/193,843
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-813-432-44

  Query Match             92.9%;  Score 1405;  DB 10;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.9e-120;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120
```

```
Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 8
US-10-352-684A-48
; Sequence 48, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-352-684A-48

  Query Match             92.9%;  Score 1405;  DB 15;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.9e-120;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120
```

```
Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275

RESULT 9
US-10-174-364-44
; Sequence 44, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson et al.
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-729CIP2
;  CURRENT APPLICATION NUMBER: US/10/174,364
;  CURRENT FILING DATE:  2002-06-17
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-174-364-44

  Query Match             92.9%;  Score 1405;  DB 15;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.9e-120;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
```

```
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 10
US-10-246-583-44
; Sequence 44, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-729CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/246,583
;  CURRENT FILING DATE:  2002-12-06
;  PRIOR APPLICATION NUMBER: 10/174,364
;  PRIOR FILING DATE: 2002-06-17
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-246-583-44

  Query Match             92.9%;  Score 1405;  DB 15;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.9e-120;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 11
US-10-287-226-94
; Sequence 94, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 94
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-287-226-94

  Query Match             92.9%;  Score 1405;  DB 15;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.9e-120;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275

RESULT 12
US-10-689-832-44
; Sequence 44, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
;  APPLICANT: Majmuder, Kamud
;  TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
;  FILE REFERENCE: 15966-729DIV1
;  CURRENT APPLICATION NUMBER: US/10/689,832
;  CURRENT FILING DATE:  2003-10-20
;  PRIOR APPLICATION NUMBER: 09/813,432
;  PRIOR FILING DATE: 2001-03-20
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-689-832-44

  Query Match             92.9%;  Score 1405;  DB 16;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 2.9e-120;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275

RESULT 13
US-10-275-505-13
; Sequence 13, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 13
;   LENGTH: 691
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7472460CD1
US-10-275-505-13

  Query Match             92.7%;  Score 1401;  DB 15;  Length 691;
  Best Local Similarity   92.0%;  Pred. No. 2e-119;
  Matches  253;  Conservative    8;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||   ||||: |||||||||||||||||||||||||||||||
Db         11 MLNLLLLALPVLASRAYAAPG--QALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 68

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         69 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 128

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| |||||||| |||
Db        129 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKHVKV 188
```

```
Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        189 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 248

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              || | ||||||||||||||||||||||||||||||
Db        249 VVRWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 283


RESULT 14
US-10-117-323-40
; Sequence 40, Application US/10117323
; Publication No. US20030054993A1
; GENERAL INFORMATION:
;  APPLICANT: Rancourt, Derrick E.
;  APPLICANT:  Rancourt, Susan L.
;  APPLICANT:  O'Sullivan, Colleen M.
;  TITLE OF INVENTION: Implantation Serine Proteinases
;  FILE REFERENCE: 033337-005
;  CURRENT APPLICATION NUMBER: US/10/117,323
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US 60/281,724
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: US 60/294,736
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: US 60/350,962
;  PRIOR FILING DATE: 2002-01-25
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Mouse
US-10-117-323-40

  Query Match             91.7%;  Score 1386.5;  DB 14;  Length 272;
  Best Local Similarity   91.9%;  Pred. No. 1.4e-118;
  Matches  251;  Conservative    8;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          3 SLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGS 62
              :||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||||
Db          1 NLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGS 60

Qy         63 LIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADI 122
              |||||||||||||:|||||||| | |||||||||||||||||||||||||||  | ||||
Db         61 LIHPQWVLTAAHCVGPDVKDLAAL-VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADI 119

Qy        123 ALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPI 182
              ||||||||||:|| |||| |||||||||||||||||||||||||||| ||||||||||||||
Db        120 ALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPI 179

Qy        183 MENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVV 242
              ||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||||
Db        180 MENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV 239

Qy        243 SWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              || ||||||||||||||||||||||||||||||
Db        240 SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 272


RESULT 15
US-09-789-210-69
; Sequence 69, Application US/09789210
; Publication No. US20040241646A1
;   GENERAL INFORMATION:
;        APPLICANT: COHEN, MAURICE
;                   COLPITTS, TRACEY L.
;                   FRIEDMAN, PAULA N.
;                   GRANADOS, EDWARD N.
;                   KLASS, MICHAEL R.
;                   RUSSELL, JOHN C.
;                   STEWART, KENT D.
;                   STROUPE, STEVEN D.
;        TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
;                            AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
;                            OF THE PROSTATE
;        NUMBER OF SEQUENCES: 76
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Abbott Laboratories
;             STREET: 100 Abbott Park Road
;             CITY: Abbott Park
;             STATE: IL
;             COUNTRY: USA
;             ZIP: 60064-3500
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/789,210
;             FILING DATE: 20-Feb-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/944,483
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Becker, Cheryl L.
;             REGISTRATION NUMBER: 35,441
;             REFERENCE/DOCKET NUMBER: 6183.US.01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 847/935-1729
;             TELEFAX: 847/938-2623
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 69:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 245 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: None
;        SEQUENCE DESCRIPTION: SEQ ID NO: 69:
US-09-789-210-69

  Query Match             90.8%;  Score 1373;  DB 11;  Length 245;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-117;
  Matches  245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 60

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 120

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 180

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        271 VPKKP 275
              |||||
Db        241 VPKKP 245


Search completed: August 27, 2005, 23:41:06
Job time : 115 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         August 27, 2005, 23:22:30 ; Search time 63 Seconds
                                            (without alignments)
                                            325.849 Million cell updates/sec

Title:          US-09-598-982C-52
Perfect score:  1512
Sequence:       1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:      513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                 1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES

                 %
Result         Query
   No.  Score  Match Length DB  ID                     Description
-----------------------------------------------------------------------------
     1   1512  100.0    275  2  US-09-016-366A-17       Sequence 17, Appl
     2   1512  100.0    275  2  US-08-978-404B-12       Sequence 12, Appl
     3   1402   92.7    273  2  US-09-016-366A-19       Sequence 19, Appl
     4   1402   92.7    273  2  US-08-978-404B-14       Sequence 14, Appl
     5   1400   92.6    274  2  US-09-016-366A-21       Sequence 21, Appl
     6   1400   92.6    274  2  US-08-978-404B-16       Sequence 16, Appl
     7   1394   92.2    267  2  US-09-016-366A-23       Sequence 23, Appl
     8   1394   92.2    267  2  US-08-978-404B-18       Sequence 18, Appl
     9   1394   92.2    267  4  US-09-917-254-101       Sequence 101, App
    10   1373   90.8    245  3  US-08-944-483-69        Sequence 69, Appl
    11   1284   84.9    245  3  US-09-079-970A-6        Sequence 6, Appli
    12   1284   84.9    245  4  US-09-601-318-1         Sequence 1, Appli
    13   1284   84.9    249  3  US-09-079-970A-5        Sequence 5, Appli
    14   1277   84.5    244  4  US-09-601-318-4         Sequence 4, Appli
    15   1277   84.5    244  4  US-09-601-318-5         Sequence 5, Appli
    16   1277   84.5    244  4  US-09-601-318-6         Sequence 6, Appli
    17   1277   84.5    244  4  US-09-601-318-7         Sequence 7, Appli
    18   1134   75.0    273  2  US-08-978-404B-3        Sequence 3, Appli
    19   1133   74.9    276  2  US-09-016-366A-15       Sequence 15, Appl
    20   1133   74.9    276  2  US-08-978-404B-21       Sequence 21, Appl
    21 1109.5   73.4    270  2  US-08-978-404B-8        Sequence 8, Appli
    22   1105   73.1    273  2  US-08-978-404B-6        Sequence 6, Appli
    23   1105   73.1    274  2  US-08-978-404B-5        Sequence 5, Appli
    24    820   54.2    190  2  US-08-845-998-4         Sequence 4, Appli
    25    820   54.2    190  3  US-09-206-537-4         Sequence 4, Appli
    26    820   54.2    190  3  US-09-430-854-4         Sequence 4, Appli
    27    816   54.0    190  2  US-08-845-998-6         Sequence 6, Appli
    28    816   54.0    190  3  US-09-206-537-6         Sequence 6, Appli
    29    816   54.0    190  3  US-09-430-854-6         Sequence 6, Appli
    30  686.5   45.4    269  2  US-08-978-404B-10       Sequence 10, Appl
    31  572.5   37.9    290  4  US-09-386-653A-7        Sequence 7, Appli
    32    555   36.7    315  4  US-09-386-653A-9        Sequence 9, Appli
    33  553.5   36.6    284  4  US-09-387-375-7         Sequence 7, Appli
    34  553.5   36.6    284  4  US-10-041-400A-7        Sequence 7, Appli
    35  553.5   36.6    284  4  US-10-042-091A-7        Sequence 7, Appli
    36    535   35.4    316  4  US-09-387-375-9         Sequence 9, Appli
    37    535   35.4    316  4  US-10-041-400A-9        Sequence 9, Appli
    38    535   35.4    316  4  US-10-042-091A-9        Sequence 9, Appli
    39  530.5   35.1    317  4  US-09-386-629-7         Sequence 7, Appli
    40  530.5   35.1    317  4  US-09-907-794A-263      Sequence 263, App
    41  530.5   35.1    317  4  US-09-905-125A-263      Sequence 263, App
    42  530.5   35.1    317  4  US-09-902-775A-263      Sequence 263, App
    43  530.5   35.1    317  4  US-09-906-700-263       Sequence 263, App
    44  530.5   35.1    317  4  US-09-903-603A-263      Sequence 263, App
    45  530.5   35.1    317  4  US-09-904-920A-263      Sequence 263, App


                                  ALIGNMENTS


RESULT 1
US-09-016-366A-17
; Sequence 17, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  275 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-016-366A-17

  Query Match             100.0%;  Score 1512;  DB 2;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-153;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 2
US-08-978-404B-12
; Sequence 12, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  275 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-12

  Query Match             100.0%;  Score 1512;  DB 2;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-153;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
```

```
Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 3
US-09-016-366A-19
; Sequence 19, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-19

  Query Match             92.7%;  Score 1402;  DB 2;  Length 273;
  Best Local Similarity   92.3%;  Pred. No. 4.4e-141;
  Matches  252;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          3 SLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGS 62
              :||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||||
Db          1 NLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGS 60

Qy         63 LIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADI 122
              |||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||||
```

```
Db          61 LIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADI 120

Qy         123 ALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPI 182
               ||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||||
Db         121 ALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPI 180

Qy         183 MENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVV 242
               ||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||||
Db         181 MENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV 240

Qy         243 SWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
               || ||||||||||||||||||||||||||||||
Db         241 SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 273

RESULT 4
US-08-978-404B-14
; Sequence 14, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-14

  Query Match             92.7%;  Score 1402;  DB 2;  Length 273;
  Best Local Similarity   92.3%;  Pred. No. 4.4e-141;
  Matches  252;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy           3 SLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGS 62
               :||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||||
Db           1 NLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGS 60

Qy          63 LIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADI 122
               |||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||||
Db          61 LIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADI 120

Qy         123 ALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPI 182
               ||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||||
Db         121 ALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPI 180

Qy         183 MENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVV 242
               ||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||||
Db         181 MENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV 240

Qy         243 SWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
               || ||||||||||||||||||||||||||||||
Db         241 SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 273

RESULT 5
US-09-016-366A-21
; Sequence 21, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-21

  Query Match             92.6%;  Score 1400;  DB 2;  Length 274;
  Best Local Similarity   92.0%;  Pred. No. 7.2e-141;
  Matches  252;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy           2 LSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGG 61
               |:||||||||||||||||||| ||||: |||||||||||||||||||||   ||||||||
Db           1 LNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGG 60

Qy          62 SLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGAD 121
               ||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | |||
Db          61 SLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAD 120

Qy         122 IALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVP 181
               |||||||||| :|| |||| |||||||||||||||||||||||||| |||||||||||||
Db         121 IALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVP 180
```

```
Qy        182 IMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGV 241
              |||||||||||||||||||||||:|||||||||::||||:||||||||||||||||||||
Db        181 IMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGV 240

Qy        242 VSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              ||| ||||||||||||||||||||||||||||||
Db        241 VSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 274

RESULT 6
US-08-978-404B-16
; Sequence 16, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-16

  Query Match             92.6%;  Score 1400;  DB 2;  Length 274;
  Best Local Similarity   92.0%;  Pred. No. 7.2e-141;
  Matches  252;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          2 LSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGG 61
              |:||||||||||||||||||| ||||: |||||||||||||||||||||   ||||||||
Db          1 LNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGG 60

Qy         62 SLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGAD 121
              ||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | |||
Db         61 SLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAD 120

Qy        122 IALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVP 181
              |||||||||| :|| |||| |||||||||||||||||||||||||| |||||||||||||
Db        121 IALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVP 180

Qy        182 IMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGV 241
              |||||||||||||||||||||||:|||||||||::||||:||||||||||||||||||||
Db        181 IMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGV 240
```

```
Qy        242 VSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              ||| ||||||||||||||||||||||||||||||
Db        241 VSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 274

RESULT 7
US-09-016-366A-23
; Sequence 23, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-23

  Query Match             92.2%;  Score 1394;  DB 2;  Length 267;
  Best Local Similarity   93.3%;  Pred. No. 3e-140;
  Matches  249;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          9 LPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQW 68
              |||||||||||||| ||||: |||||||||||||||||||||||||||||||||||||||
Db          1 LPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQW 60

Qy         69 VLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELE 128
              |||||||:|||||||| |||||||||||||||||||||||||||||  | ||||||||||
Db         61 VLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE 120

Qy        129 EPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHIC 188
              ||| :|| |||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHIC 180

Qy        189 DAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGC 248
              ||||||||||||||||:|||||||||::||||:||||||||||||||||||||||| |||
Db        181 DAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC 240

Qy        249 AQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||
```

```
Db        241 AQPNRPGIYTRVTYYLDWIHHYVPKKP 267


RESULT 8
US-08-978-404B-18
; Sequence 18, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-18

  Query Match             92.2%;  Score 1394;  DB 2;  Length 267;
  Best Local Similarity   93.3%;  Pred. No. 3e-140;
  Matches  249;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          9 LPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQW 68
              |||||||||||||| ||||: |||||||||||||||||||||||||||||||||||||||
Db          1 LPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQW 60

Qy         69 VLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELE 128
              |||||||:|||||||| |||||||||||||||||||||||||||||  | ||||||||||
Db         61 VLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE 120

Qy        129 EPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHIC 188
              ||| :|| |||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHIC 180

Qy        189 DAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGC 248
              ||||||||||||||||:|||||||||::||||:||||||||||||||||||||||| |||
Db        181 DAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC 240

Qy        249 AQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||
Db        241 AQPNRPGIYTRVTYYLDWIHHYVPKKP 267
```

```
RESULT 9
US-09-917-254-101
; Sequence 101, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Baak, Jan
;  TITLE OF INVENTION: Prognostic Classification of Breast Cancer
;  FILE REFERENCE: B0801/7224(JRV)
;  CURRENT APPLICATION NUMBER: US/09/917,254
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/222,093
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 101
;   LENGTH: 267
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-917-254-101

  Query Match             92.2%;  Score 1394;  DB 4;  Length 267;
  Best Local Similarity   93.3%;  Pred. No. 3e-140;
  Matches  249;  Conservative    7;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          9 LPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQW 68
              |||||||||||||| ||||: |||||||||||||||||||||||||||||||||||||||
Db          1 LPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQW 60

Qy         69 VLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELE 128
              |||||||:|||||||| |||||||||||||||||||||||||||||  | ||||||||||
Db         61 VLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE 120

Qy        129 EPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHIC 188
              ||| :|| |||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHIC 180

Qy        189 DAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGC 248
              ||||||||||||||||:|||||||||::||||:||||||||||||||||||||||| |||
Db        181 DAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC 240

Qy        249 AQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||
Db        241 AQPNRPGIYTRVTYYLDWIHHYVPKKP 267


RESULT 10
US-08-944-483-69
; Sequence 69, Application US/08944483
; Patent No. 6232456
;  GENERAL INFORMATION:
;    APPLICANT:  COHEN, MAURICE
;    APPLICANT:  COLPITTS, TRACEY L.
;    APPLICANT:  FRIEDMAN, PAULA N.
;    APPLICANT:  GRANADOS, EDWARD N.
;    APPLICANT:  KLASS, MICHAEL R.
;    APPLICANT:  RUSSELL, JOHN C.
;    APPLICANT:  STEWART, KENT D.
;    APPLICANT:  STROUPE, STEVEN D.
;    TITLE OF INVENTION:  NOVEL SERINE PROTEASE REAGENTS
;    TITLE OF INVENTION:  AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
;    TITLE OF INVENTION:  OF THE PROSTATE
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
```

```
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,483
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6183.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  245 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6232456e
US-08-944-483-69

  Query Match             90.8%;  Score 1373;  DB 3;  Length 245;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-138;
  Matches  245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 60

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 120

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 180

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        271 VPKKP 275
              |||||
Db        241 VPKKP 245

RESULT 11
US-09-079-970A-6
; Sequence 6, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  245 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-079-970A-6

  Query Match             84.9%;  Score 1284;  DB 3;  Length 245;
  Best Local Similarity   92.7%;  Pred. No. 1.4e-128;
  Matches  227;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||   ||||||||||||||||||||||:|||||||| |||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        271 VPKKP 275
              |||||
Db        241 VPKKP 245

RESULT 12
US-09-601-318-1
; Sequence 1, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
```

```
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 245
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-1

  Query Match             84.9%;  Score 1284;  DB 4;  Length 245;
  Best Local Similarity   92.7%;  Pred. No. 1.4e-128;
  Matches  227;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||   ||||||||||||||||||||||:|||||||| |||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        271 VPKKP 275
              |||||
Db        241 VPKKP 245


RESULT 13
US-09-079-970A-5
; Sequence 5, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  249 amino acids
```

```
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-079-970A-5

  Query Match             84.9%;  Score 1284;  DB 3;  Length 249;
  Best Local Similarity   92.7%;  Pred. No. 1.5e-128;
  Matches  227;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||   ||||||||||||||||||||||:|||||||| |||||
Db          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 64

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244

Qy        271 VPKKP 275
              |||||
Db        245 VPKKP 249


RESULT 14
US-09-601-318-4
; Sequence 4, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-4

  Query Match             84.5%;  Score 1277;  DB 4;  Length 244;
  Best Local Similarity   92.6%;  Pred. No. 7.9e-128;
  Matches  226;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;
```

```
Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||   ||||||||||||||||||||||:|||||||| |||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        271 VPKK 274
              ||||
Db        241 VPKK 244


RESULT 15
US-09-601-318-5
; Sequence 5, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-5

  Query Match             84.5%;  Score 1277;  DB 4;  Length 244;
  Best Local Similarity   92.6%;  Pred. No. 7.9e-128;
  Matches  226;  Conservative    6;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90
              ||||||||||||||||||||   ||||||||||||||||||||||:|||||||| |||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        271 VPKK 274
              ||||
Db        241 VPKK 244

Search completed: August 27, 2005, 23:31:26
Job time : 64 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          August 27, 2005, 23:12:00 ; Search time 124 Seconds
                                          (without alignments)
                                          857.735 Million cell updates/sec

Title:           US-09-598-982C-52
Perfect score:   1512
Sequence:        1 MLSLLLLALPVLASRAYAAP..........IYTRVTYYLDWIHHYVPKKP 275

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:      2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        A_Geneseq_16Dec04:*
                  1:  geneseqp1980s:*
                  2:  geneseqp1990s:*
                  3:  geneseqp2000s:*
                  4:  geneseqp2001s:*
                  5:  geneseqp2002s:*
                  6:  geneseqp2003as:*
                  7:  geneseqp2003bs:*
                  8:  geneseqp2004s:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                    %
Result              Query
  No.    Score    Match Length DB  ID                  Description
 -------------------------------------------------------------------------
     1    1512    100.0    275  2  AAW64237            Aaw64237 Human mas
     2    1512    100.0    275  2  AAW63173            Aaw63173 Human mas
     3    1512    100.0    275  7  ADK52588            Adk52588 Hematolog
     4    1512    100.0    275  8  ADP12448            Adp12448 Protein e
     5    1512    100.0    275  8  ADP56070            Adp56070 Human PRO
     6    1505     99.5    275  7  ADE62892            Ade62892 Human Pro
     7    1430     94.6    275  7  ADE56171            Ade56171 Human Pro
     8    1411     93.3    275  7  ADJ94864            Adj94864 Novel NOV
     9    1411     93.3    275  8  ADQ19871            Adq19871 Human sof
    10    1405     92.9    275  7  ADJ94866            Adj94866 Novel NOV
    11    1405     92.9    275  7  ADK52590            Adk52590 Hematolog
    12    1402     92.7    273  2  AAW64238            Aaw64238 Human mas
    13    1402     92.7    273  2  AAW63174            Aaw63174 Human mas
    14    1401     92.7    691  5  AAE14348            Aae14348 Human pro
    15    1400     92.6    274  2  AAW64240            Aaw64240 Human mas
    16    1400     92.6    274  2  AAW63175            Aaw63175 Human mas
    17    1394     92.2    267  2  AAW64241            Aaw64241 Human mas
    18    1394     92.2    267  2  AAW63176            Aaw63176 Human mas
    19    1394     92.2    267  5  AAU84360            Aau84360 Protein T
    20    1382     91.4    275  7  ADJ94862            Adj94862 Novel NOV
    21    1360     89.9    264  7  ADG42743            Adg42743 Human bet
    22    1360     89.9    264  8  ADM76650            Adm76650 Human NOV
    23    1349     89.2    266  7  ADG42723            Adg42723 Human mas
    24    1349     89.2    266  8  ADM76630            Adm76630 Human NOV
    25  1346.5     89.1    333  8  ADJ72001            Adj72001 Human PMM
    26    1324     87.6    256  7  ADG42731            Adg42731 Human bet
    27    1324     87.6    256  8  ADM76638            Adm76638 Human NOV
    28    1323     87.5    264  7  ADJ55812            Adj55812 Peptide h
    29    1316     87.0    254  7  ADJ55782            Adj55782 Peptide h
    30    1313     86.8    254  7  ADG42713            Adg42713 Human bet
    31    1313     86.8    254  8  ADM76620            Adm76620 Human NOV
    32    1307     86.4    256  7  ADJ55800            Adj55800 Peptide h
    33  1298.5     85.9    275  7  ADJ94860            Adj94860 Novel NOV
    34    1290     85.3    245  5  AAU12005            Aau12005 Human mat
    35    1290     85.3    249  5  AAU12006            Aau12006 Human bet
    36    1287     85.1    245  5  AAU12023            Aau12023 Human bet
    37    1287     85.1    245  5  AAU12024            Aau12024 Human bet
    38    1287     85.1    249  5  AAU12020            Aau12020 Human bet
    39    1287     85.1    249  5  AAU12019            Aau12019 Human bet
    40    1284     84.9    245  5  AAU12008            Aau12008 Recombina
    41    1284     84.9    249  3  AAY55011            Aay55011 Human bet
    42    1284     84.9    249  5  AAU12007            Aau12007 Recombina
    43    1282     84.8    245  5  AAU12022            Aau12022 Human bet
    44    1282     84.8    249  5  AAU12018            Aau12018 Human bet
    45    1281     84.7    245  5  AAU12016            Aau12016 Human bet


                              ALIGNMENTS


RESULT 1
AAW64237
ID   AAW64237 standard; protein; 275 AA.
XX
AC   AAW64237;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human mast cell tryptase alpha.
XX
KW   Mast cell tryptase alpha; human; MCP-7; mast cell protease 7; blood clot;
KW   anticoagulant; myocardial infarction; reocclusion; thromboembolism;
KW   cerebral embolism; thrombosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .20
FT                   /label= Sig_peptide
FT   Peptide         21. .30
FT                   /label= Pro_peptide
FT   Protein         31. .275
FT                   /label= Mat_protein
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
DR   N-PSDB; AAV44328.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Disclosure; Page 63-64; 92pp; English.
XX
CC   This is the deduced amino acid sequence of human mast cell tryptase alpha
CC   (see also AAV44328). The invention provides: compositions comprising an
CC   isolated tryptase-7 that may include chimeric proteins that contain (a) a
```

```
CC   human tryptase for all but the active site region and (b) the substrate-
CC   binding pocket of mouse tryptase-7 or its homologues (see AAW64233-39); a
CC   method for treating a blood clot by administering a nucleic acid molecule
CC   that codes for a tryptase-7, or an expression product, to decrease
CC   fibrinogen activity; a nucleic acid encoding a serine protease (SP); and
CC   a method of producing a mature SP by expressing the inactive zymogen in a
CC   host cell, and cleaving the enterokinase susceptibility domain. The
CC   tryptase-7 polypeptides can be used to treat disorders mediated by
CC   undesirable thrombus clot formation such as myocardial infarction and
CC   reocclusion following angioplasty of blood clots associated with
CC   pulmonary thromboembolism, deep vein thrombosis, cerebral embolism, renal
CC   vein and peripheral arterial thrombosis. They are also useful for all
CC   surgical procedures that require decreased blood clots
XX
SQ   Sequence 275 AA;

  Query Match             100.0%;  Score 1512;  DB 2;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-131;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 2
AAW63173
ID   AAW63173 standard; protein; 275 AA.
XX
AC   AAW63173;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human mast cell tryptase alpha polypeptide.
XX
KW   Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
KW   tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
KW   antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
KW   hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
KW   inflammatory skin condition; human; mast cell tryptase alpha.
XX
OS   Homo sapiens.
XX
PN   WO9833812-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US001865.
XX
PR   05-FEB-1997;   97US-0037090P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL,  Huang C;
XX
DR   WPI; 1998-437390/37.
DR   N-PSDB; AAV42710.
```

```
XX
PT   Tryptase-6 complex inhibitory peptides - used to treat mast cell-mediated
PT   inflammatory disorders e.g. asthma.
XX
PS   Disclosure; Page 45-46; 69pp; English.
XX
CC   This reoresents the human mast cell tryptase alpha protein which is a
CC   homologue of the mouse mast cell protease (mMCP-6) zymogen. The invention
CC   provides sequences shown in AAW63160 to AAW63169 that are inhibitors of
CC   mMCP-6. These peptides which are tryptase-6 complex inhibitors, can be
CC   used for treating a mast cell-mediated inflammatory disorder. The
CC   inhibitors can be used to treat inflammatory disorders including asthma,
CC   allergic rhinitis, urticaria and antioedema, eczematous dermatitis
CC   (atopic dermatitis), anaphylaxis, hyperproliferative skin disease, peptic
CC   ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory
CC   skin conditions
XX
SQ   Sequence 275 AA;

  Query Match             100.0%;  Score 1512;  DB 2;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-131;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 3
ADK52588
ID   ADK52588 standard; protein; 275 AA.
XX
AC   ADK52588;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Hematological disorder associated Gene ID 1847 encoded protein.
XX
KW   cytostatic; antianemic; antisickling; virucide; hemostatic; nephrotropic;
KW   cytostatic; thrombolytic; antiparasitic; gene therapy;
KW   hematologic disorder; cancer; Sickle Cell Anemia;
KW   Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
KW   Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
KW   transfusion reaction; Erythroblastosis; mechanical trauma;
KW   micro-angiopathic hemolytic anemia; parasite infection.
XX
OS   Homo sapiens.
XX
PN   WO2003065871-A2.
XX
PD   14-AUG-2003.
XX
PF   28-JAN-2003; 2003WO-US002484.
XX
PR   04-FEB-2002; 2002US-0354333P.
PR   28-FEB-2002; 2002US-0360258P.
PR   15-MAR-2002; 2002US-0364476P.
```

```
PR   26-APR-2002; 2002US-0375626P.
PR   06-JUN-2002; 2002US-0386494P.
PR   24-JUN-2002; 2002US-0390965P.
PR   28-JUN-2002; 2002US-0392480P.
PR   03-JUL-2002; 2002US-0394128P.
PR   31-JUL-2002; 2002US-0399783P.
PR   13-AUG-2002; 2002US-0403221P.
PR   30-AUG-2002; 2002US-0407045P.
PR   25-NOV-2002; 2002US-0429048P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Carroll JM,  Healy A,  Weich NS,  Kelly LM;
XX
DR   WPI; 2003-731464/69.
DR   N-PSDB; ADK52587.
XX
PT   Identifying a compound capable of treating a hematologic disorder (e.g.
PT   anemia or leukemia) comprises assaying the ability of the compound to
PT   modulate the expression or activity of e.g. 131,148, 199 or 12303
PT   polypeptide or nucleic acid.
XX
PS   Disclosure; SEQ ID NO 46; 232pp; English.
XX
CC   The invention relates to a method of identifying a compound capable of
CC   treating a hematologic disorder comprises assaying the ability of the
CC   compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
CC   194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
CC   13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
CC   acid expression or polypeptide activity, thus, identifying a compound
CC   capable of treating a hematologic disorder. The methods are useful in
CC   diagnosing, preventing and treating hematological disorders, such as
CC   cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC   Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders
CC   associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC   antibody-mediated disorders such as transfusion reactions and
CC   Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC   angiopathic hemolytic anemias, infections by parasites or chemical
CC   injuries. The methods may also be used for identifying compounds that
CC   modulate hematological disorders. This sequence corresponds to the
CC   protein encoded by one of the genes modulated by the compounds.
XX
SQ   Sequence 275 AA;

  Query Match             100.0%;  Score 1512;  DB 7;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-131;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 4
ADP12448
ID   ADP12448 standard; protein; 275 AA.
XX
```

```
AC   ADP12448;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Protein encoded by mRNA of the invention #58.
XX
KW   transplant rejection; immune system; rheumatoid arthritis; lupus;
KW   inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS   Homo sapiens.
XX
PN   WO2004042346-A2.
XX
PD   21-MAY-2004.
XX
PF   24-APR-2003; 2003WO-US012946.
XX
PR   24-APR-2002; 2002US-00131831.
PR   20-DEC-2002; 2002US-00325899.
XX
PA   (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
PI   Wohlgemuth J,  Fry K,  Woodward R,  Ly N,  Prentice J,  Morris M;
PI   Rosenberg S;
XX
DR   WPI; 2004-400724/37.
XX
PT   Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT   pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT   rejection, in an individual, comprises detecting the expression level of
PT   the genes.
XX
PS   Claim 65; SEQ ID NO 2457; 1762pp; English.
XX
CC   The present invention relates to diagnosing or monitoring transplant
CC   rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC   comprises detecting the expression level of one or more genes. The
CC   methods, system and kits are useful in diagnosing or monitoring
CC   transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC   islet, lung, bone marrow or stem cell transplant rejection,
CC   xenotransplant rejection or mechanical organ replacement rejection, in an
CC   individual. The method is also useful in assessing the immune status of
CC   an individual. The methods are also useful in diagnosing and monitoring
CC   diseases that involve the immune system, e.g. rheumatoid arthritis,
CC   lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC   viral, bacterial or fungal infection. The present sequence represents a
CC   protein that is encoded by the mRNA of the invention.
XX
SQ   Sequence 275 AA;

  Query Match             100.0%;  Score 1512;  DB 8;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-131;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
```

```
RESULT 5
ADP56070
ID   ADP56070 standard; protein; 275 AA.
XX
AC   ADP56070;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human PRO protein sequence SEQ ID NO:2046.
XX
KW   human; PRO; immune related disease; inflammatory immune response;
KW   immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW   antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW   antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW   haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW   nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW   virucide; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004039956-A2.
XX
PD   13-MAY-2004.
XX
PF   28-OCT-2003; 2003WO-US034381.
XX
PR   29-OCT-2002; 2002US-0422472P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Aggarwal S,  Clark H,  Gurney AL,  Schoenfeld J,  Williams PM;
PI   Wood WI,  Wu TD;
XX
DR   WPI; 2004-376182/35.
DR   N-PSDB; ADP56069.
XX
PT   New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT   and treating an immune related disease, e.g. systemic lupus
PT   erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT   stimulating an immune response.
XX
PS   Claim 1; SEQ ID NO 2046; 3009pp; English.
XX
CC   The present invention describes an isolated PRO nucleic acid (I). Also
CC   described: (1) a vector comprising (I); (2) a host cell comprising the
CC   vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC   isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC   polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC   antibody which specifically binds to a polypeptide of (4); (7) a
CC   composition of matter comprising a polypeptide of (4), an agonist or
CC   antagonist of the polypeptide or an antibody that binds to the
CC   polypeptide in combination with a carrier; (8) an article of manufacture
CC   comprising a container, a label on the container and a composition of
CC   matter of (7); (9) a method of treating an immune related disease in a
CC   mammal; (10) a method for determining the presence of a PRO polypeptide
CC   in a sample suspected of having the polypeptide; (11) a method of
CC   diagnosing an immune related disease or an inflammatory immune response
CC   in mammal; (12) a method of identifying a compound that inhibits or
CC   mimics the activity of or expression of a gene encoding a PRO polypeptide
CC   ; and (13) a method of stimulating the immune response in a mammal. The
CC   PRO sequences have antiallergic, antianaemic, antiarthritic,
CC   antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC   antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC   haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC   nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC   virucide activities, and can be used in gene therapy. The nucleic acid
CC   (I) and the encoded polypeptides, compositions, kits and methods are
CC   useful in diagnosing and treating an immune related disease and in
CC   stimulating an immune response. The present sequence represents a human
CC   PRO protein from the present invention.
XX
SQ   Sequence 275 AA;

  Query Match             100.0%;  Score 1512;  DB 8;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-131;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 6
ADE62892
ID   ADE62892 standard; protein; 275 AA.
XX
AC   ADE62892;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein P15157, SEQ ID NO 8826.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P15157.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
```

```
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 275 AA;

  Query Match             99.5%;  Score 1505;  DB 7;  Length 275;
  Best Local Similarity   99.3%;  Pred. No. 5.5e-131;
  Matches  273;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||||||||||||::||||||||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNTRRDSCKGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||||||||||||||||||||||||||||||||||
Db        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 7
ADE56171
ID   ADE56171 standard; protein; 275 AA.
XX
AC   ADE56171;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein XP_018104, SEQ ID NO 2020.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; XP_018104.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a human protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 275 AA;

  Query Match             94.6%;  Score 1430;  DB 7;  Length 275;
  Best Local Similarity   93.5%;  Pred. No. 4.9e-124;
  Matches  257;  Conservative    8;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||||||||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 8
ADJ94864
ID   ADJ94864 standard; protein; 275 AA.
XX
```

```
AC   ADJ94864;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Novel NOVX protein sequence #46.
XX
KW   antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
KW   anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
KW   neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;
KW   antiarthritic; antiinflammatory; dermatological; antiasthmatica;
KW   antilipemic; gene therapy; metabolic disorder; diabetes; obesity;
KW   infectious disease; anorexia; cancer; cardiovascular disease;
KW   hypertension; atherosclerosis; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
KW   osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
KW   asthma; dyslipidemia; neurogenesis; cell differentiation;
KW   cell proliferation; hematopoiesis; wound healing; angiogenesis;
KW   chromosome mapping; tissue typing; pharmacogenomic.
XX
OS   Homo sapiens.
XX
PN   WO2003040325-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035464.
XX
PR   05-NOV-2001; 2001US-0338626P.
PR   06-NOV-2001; 2001US-0333072P.
PR   09-NOV-2001; 2001US-0348283P.
PR   15-NOV-2001; 2001US-0335610P.
PR   16-NOV-2001; 2001US-0338543P.
PR   20-NOV-2001; 2001US-0331630P.
PR   20-NOV-2001; 2001US-0331641P.
PR   21-NOV-2001; 2001US-0332152P.
PR   27-NOV-2001; 2001US-0333461P.
PR   28-NOV-2001; 2001US-0333912P.
PR   28-NOV-2001; 2001US-0334027P.
PR   29-NOV-2001; 2001US-0334300P.
PR   30-NOV-2001; 2001US-0334421P.
PR   30-NOV-2001; 2001US-0334526P.
PR   04-DEC-2001; 2001US-0336576P.
PR   04-DEC-2001; 2001US-0336664P.
PR   07-DEC-2001; 2001US-0338314P.
PR   07-DEC-2001; 2001US-0338390P.
PR   10-DEC-2001; 2001US-0339006P.
PR   10-DEC-2001; 2001US-0339008P.
PR   11-DEC-2001; 2001US-0339286P.
PR   01-FEB-2002; 2002US-0353280P.
PR   01-FEB-2002; 2002US-0353288P.
PR   04-FEB-2002; 2002US-0354392P.
PR   04-FEB-2002; 2002US-0354393P.
PR   04-FEB-2002; 2002US-0354409P.
PR   27-FEB-2002; 2002US-0359944P.
PR   27-FEB-2002; 2002US-0360148P.
PR   05-MAR-2002; 2002US-0361790P.
PR   05-MAR-2002; 2002US-0361833P.
PR   05-MAR-2002; 2002US-0361925P.
PR   05-MAR-2002; 2002US-0362230P.
PR   05-MAR-2002; 2002US-0362625P.
PR   13-MAR-2002; 2002US-0364000P.
PR   13-MAR-2002; 2002US-0364181P.
PR   13-MAR-2002; 2002US-0364182P.
PR   13-MAR-2002; 2002US-0364197P.
PR   13-MAR-2002; 2002US-0364227P.
PR   17-MAY-2002; 2002US-0381621P.
PR   28-MAY-2002; 2002US-0383675P.
PR   17-JUL-2002; 2002US-0396703P.
PR   06-AUG-2002; 2002US-0401552P.
PR   07-AUG-2002; 2002US-0401594P.
PR   07-AUG-2002; 2002US-0401787P.
PR   15-AUG-2002; 2002US-0403619P.
PR   20-AUG-2002; 2002US-0404821P.
PR   23-AUG-2002; 2002US-0405368P.
PR   23-AUG-2002; 2002US-0405402P.
PR   23-AUG-2002; 2002US-0405496P.
PR   23-AUG-2002; 2002US-0405631P.
PR   26-AUG-2002; 2002US-0406125P.
PR   04-NOV-2002; 2002US-00287226.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Agee ML,  Alsobrook JP,  Berghs C,  Boldog FL,  Burgess CE,  Chant JS;
PI   Chaudhuri A,  Dipippo VA,  Edinger SR,  Eisen A,  Ellerman K;
PI   Gangolli EA,  Gorman L,  Gerlach VL,  Ji W,  Kekuda R,  Khramtsov NV;
PI   Li L,  Malyankar UM,  Macdougall JR,  Mezes PS,  Miller CE,  Millet I;
PI   Ooi CE,  Ort T,  Padigaru M,  Patturajan M,  Rastelli L,  Rieger DK;
PI   Rothenberg ME,  Shenoy SG,  Spaderna SK,  Spytek KA,  Taupier RJ;
PI   Vernet CAM,  Zerhusen BD,  Zhong M;
XX
DR   WPI; 2003-441551/41.
DR   N-PSDB; ADJ94863.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 1; SEQ ID NO 92; 800pp; English.
XX
CC   The invention relates to novel isolated polypeptides, mature forms of
CC   these, or a sequence that is at least 95 % identical to, or having one or
CC   more conservative amino acid substitutions in the polypeptides. The
CC   polypeptides, nucleic acid molecules and antibodies are useful in the
CC   manufacture of a medicament for treating a syndrome associated with a
CC   human disease, preferably a NOVX-associated disorder. The nucleic acid
CC   molecules, polypeptides and antibodies are useful for treating,
CC   preventing or diagnosing diseases such metabolic disorders, diabetes,
CC   obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
CC   protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
CC   atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
CC   Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
CC   hematopoietic disorders, inflammatory skin disorders, asthma, and various
CC   dyslipidemias. The nucleic acids and polypeptides may also be used as
CC   targets for the identification of small molecules that modulate or
CC   inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
CC   hematopoiesis, wound healing and angiogenesis, in gene therapy, in
CC   generation of antibodies that bind immunospecifically to NOVX substances
CC   for use in therapeutic or diagnostic methods. The nucleic acids are
CC   further used as hybridization probes, in chromosome mapping, tissue
CC   typing, preventive medicine, and pharmacogenomics. This sequence
CC   corresponds to one of the NOVX polypeptides of the invention.
XX
SQ   Sequence 275 AA;

  Query Match             93.3%;  Score 1411;  DB 7;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 2.9e-122;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
```

```
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 9
ADQ19871
ID   ADQ19871 standard; protein; 275 AA.
XX
AC   ADQ19871;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated protein - SEQ ID 2690.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 2690; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 275 AA;

  Query Match             93.3%;  Score 1411;  DB 8;  Length 275;
  Best Local Similarity   92.4%;  Pred. No. 2.9e-122;
  Matches  254;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
```

```
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 10
ADJ94866
ID   ADJ94866 standard; protein; 275 AA.
XX
AC   ADJ94866;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Novel NOVX protein sequence #47.
XX
KW   antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
KW   anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
KW   neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;
KW   antiarthritic; antiinflammatory; dermatological; antiasthmatica;
KW   antilipemic; gene therapy; metabolic disorder; diabetes; obesity;
KW   infectious disease; anorexia; cancer; cardiovascular disease;
KW   hypertension; atherosclerosis; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
KW   osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
KW   asthma; dyslipidemia; neurogenesis; cell differentiation;
KW   cell proliferation; hematopoiesis; wound healing; angiogenesis;
KW   chromosome mapping; tissue typing; pharmacogenomic.
XX
OS   Homo sapiens.
XX
PN   WO2003040325-A2.
XX
PD   15-MAY-2003.
XX
PF   05-NOV-2002; 2002WO-US035464.
XX
PR   05-NOV-2001; 2001US-0338626P.
PR   06-NOV-2001; 2001US-0333072P.
PR   09-NOV-2001; 2001US-0348283P.
PR   15-NOV-2001; 2001US-0335610P.
PR   16-NOV-2001; 2001US-0338543P.
PR   20-NOV-2001; 2001US-0331630P.
PR   20-NOV-2001; 2001US-0331641P.
PR   21-NOV-2001; 2001US-0332152P.
PR   27-NOV-2001; 2001US-0333461P.
PR   28-NOV-2001; 2001US-0333912P.
PR   28-NOV-2001; 2001US-0334027P.
PR   29-NOV-2001; 2001US-0334300P.
PR   30-NOV-2001; 2001US-0334421P.
PR   30-NOV-2001; 2001US-0334526P.
PR   04-DEC-2001; 2001US-0336576P.
PR   04-DEC-2001; 2001US-0336664P.
PR   07-DEC-2001; 2001US-0338314P.
PR   07-DEC-2001; 2001US-0338390P.
PR   10-DEC-2001; 2001US-0339006P.
PR   10-DEC-2001; 2001US-0339008P.
PR   11-DEC-2001; 2001US-0339286P.
PR   01-FEB-2002; 2002US-0353280P.
PR   01-FEB-2002; 2002US-0353288P.
PR   04-FEB-2002; 2002US-0354392P.
PR   04-FEB-2002; 2002US-0354393P.
PR   04-FEB-2002; 2002US-0354409P.
PR   27-FEB-2002; 2002US-0359944P.
PR   27-FEB-2002; 2002US-0360148P.
PR   05-MAR-2002; 2002US-0361790P.
PR   05-MAR-2002; 2002US-0361833P.
PR   05-MAR-2002; 2002US-0361925P.
PR   05-MAR-2002; 2002US-0362230P.
PR   05-MAR-2002; 2002US-0362625P.
PR   13-MAR-2002; 2002US-0364000P.
PR   13-MAR-2002; 2002US-0364181P.
```

```
PR   13-MAR-2002; 2002US-0364182P.
PR   13-MAR-2002; 2002US-0364197P.
PR   13-MAR-2002; 2002US-0364227P.
PR   17-MAY-2002; 2002US-0381621P.
PR   28-MAY-2002; 2002US-0383675P.
PR   17-JUL-2002; 2002US-0396703P.
PR   06-AUG-2002; 2002US-0401552P.
PR   07-AUG-2002; 2002US-0401594P.
PR   07-AUG-2002; 2002US-0401787P.
PR   15-AUG-2002; 2002US-0403619P.
PR   20-AUG-2002; 2002US-0404821P.
PR   23-AUG-2002; 2002US-0405368P.
PR   23-AUG-2002; 2002US-0405402P.
PR   23-AUG-2002; 2002US-0405496P.
PR   23-AUG-2002; 2002US-0405631P.
PR   26-AUG-2002; 2002US-0406125P.
PR   04-NOV-2002; 2002US-00287226.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Agee ML,  Alsobrook JP,  Berghs C,  Boldog FL,  Burgess CE,  Chant JS;
PI   Chaudhuri A,  Dipippo VA,  Edinger SR,  Eisen A,  Ellerman K;
PI   Gangolli EA,  Gorman L,  Gerlach VL,  Ji W,  Kekuda R,  Khramtsov NV;
PI   Li L,  Malyankar UM,  Macdougall JR,  Mezes PS,  Miller CE,  Millet I;
PI   Ooi CE,  Ort T,  Padigaru M,  Patturajan M,  Rastelli L,  Rieger DK;
PI   Rothenberg ME,  Shenoy SG,  Spaderna SK,  Spytek KA,  Taupier RJ;
PI   Vernet CAM,  Zerhusen BD,  Zhong M;
XX
DR   WPI; 2003-441551/41.
DR   N-PSDB; ADJ94865.
XX
PT   New isolated NOVX polypeptides and polynucleotides, useful for
PT   preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT   osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT   asthma, or infections.
XX
PS   Claim 1; SEQ ID NO 94; 800pp; English.
XX
CC   The invention relates to novel isolated polypeptides, mature forms of
CC   these, or a sequence that is at least 95 % identical to, or having one or
CC   more conservative amino acid substitutions in the polypeptides. The
CC   polypeptides, nucleic acid molecules and antibodies are useful in the
CC   manufacture of a medicament for treating a syndrome associated with a
CC   human disease, preferably a NOVX-associated disorder. The nucleic acid
CC   molecules, polypeptides and antibodies are useful for treating,
CC   preventing or diagnosing diseases such metabolic disorders, diabetes,
CC   obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
CC   protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
CC   atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
CC   Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
CC   hematopoietic disorders, inflammatory skin disorders, asthma, and various
CC   dyslipidemias. The nucleic acids and polypeptides may also be used as
CC   targets for the identification of small molecules that modulate or
CC   inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
CC   hematopoiesis, wound healing and angiogenesis, in gene therapy, in
CC   generation of antibodies that bind immunospecifically to NOVX substances
CC   for use in therapeutic or diagnostic methods. The nucleic acids are
CC   further used as hybridization probes, in chromosome mapping, tissue
CC   typing, preventive medicine, and pharmacogenomics. This sequence
CC   corresponds to one of the NOVX polypeptides of the invention.
XX
SQ   Sequence 275 AA;

  Query Match             92.9%;  Score 1405;  DB 7;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 1e-121;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db          1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||| :|| |||| |||||||||||||||||||||||||| ||||||||||||
Db        121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              |||| ||||||||||||||||||||||||||||||
Db        241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 11
ADK52590
ID   ADK52590 standard; protein; 275 AA.
XX
AC   ADK52590;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Hematological disorder associated Gene ID 1849 encoded protein.
XX
KW   cytostatic; antianemic; antisickling; virucide; hemostatic; nephrotropic;
KW   cytostatic; thrombolytic; antiparasitic; gene therapy;
KW   hematologic disorder; cancer; Sickle Cell Anemia;
KW   Infectious Mononucleosis; Leukemia; Polycythemia Vera; Lymphoma;
KW   Retinoblastoma; Hemophilia; Thrombosis; Herpes; Thalassemia;
KW   transfusion reaction; Erythroblastosis; mechanical trauma;
KW   micro-angiopathic hemolytic anemia; parasite infection.
XX
OS   Homo sapiens.
XX
PN   WO2003065871-A2.
XX
PD   14-AUG-2003.
XX
PF   28-JAN-2003; 2003WO-US002484.
XX
PR   04-FEB-2002; 2002US-0354333P.
PR   28-FEB-2002; 2002US-0360258P.
PR   15-MAR-2002; 2002US-0364476P.
PR   26-APR-2002; 2002US-0375626P.
PR   06-JUN-2002; 2002US-0386494P.
PR   24-JUN-2002; 2002US-0390965P.
PR   28-JUN-2002; 2002US-0392480P.
PR   03-JUL-2002; 2002US-0394128P.
PR   31-JUL-2002; 2002US-0399783P.
PR   13-AUG-2002; 2002US-0403221P.
PR   30-AUG-2002; 2002US-0407045P.
PR   25-NOV-2002; 2002US-0429048P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Carroll JM,  Healy A,  Weich NS,  Kelly LM;
XX
DR   WPI; 2003-731464/69.
DR   N-PSDB; ADK52589.
XX
PT   Identifying a compound capable of treating a hematologic disorder (e.g.
PT   anemia or leukemia) comprises assaying the ability of the compound to
PT   modulate the expression or activity of e.g. 131,148, 199 or 12303
PT   polypeptide or nucleic acid.
XX
PS   Disclosure; SEQ ID NO 48; 232pp; English.
XX
CC   The invention relates to a method of identifying a compound capable of
CC   treating a hematologic disorder comprises assaying the ability of the
CC   compound to modulate 131,148, 199, 12303, 13906, 15513, 17822, 302, 5677,
CC   194, 14393, 28059, 7366, 12212, 1981, 261, 12416, 270, 1410, 137, 1871,
CC   13051, 1847, 1849, 15402, 340, 10217, 837, 1761, 8990 or 13249 nucleic
```

```
CC   acid expression or polypeptide activity, thus, identifying a compound
CC   capable of treating a hematologic disorder. The methods are useful in
CC   diagnosing, preventing and treating hematological disorders, such as
CC   cancer, Sickle Cell Anemia, Infectious Mononucleosis, Leukemia,
CC   Polycythemia Vera, Lymphoma, Retinoblastoma, Hemophilia, disorders
CC   associated with an increased risk of Thrombosis, Herpes, Thalassemia,
CC   antibody-mediated disorders such as transfusion reactions and
CC   Erythroblastosis, mechanical trauma to red blood cells such as micro-
CC   angiopathic hemolytic anemias, infections by parasites or chemical
CC   injuries. The methods may also be used for identifying compounds that
CC   modulate hematological disorders. This sequence corresponds to the
CC   protein encoded by one of the genes modulated by the compounds.
XX
SQ   Sequence 275 AA;

  Query Match             92.9%;  Score 1405;  DB 7;  Length 275;
  Best Local Similarity   92.0%;  Pred. No. 1e-121;
  Matches  253;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy           1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
               ||:||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||
Db           1 MLNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCG 60

Qy          61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
               |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db          61 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 120

Qy         121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
               ||||||||||| :|| |||| |||||||||||||||||||||||||||| ||||||||||||
Db         121 DIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKV 180

Qy         181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
               ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db         181 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 240

Qy         241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
               |||| ||||||||||||||||||||||||||||||
Db         241 VVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275


RESULT 12
AAW64238
ID   AAW64238 standard; protein; 273 AA.
XX
AC   AAW64238;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human mast cell tryptase I.
XX
KW   Mast cell tryptase I; human; MCP-7; mast cell protease 7; blood clot;
KW   anticoagulant; myocardial infarction; reocclusion; thromboembolism;
KW   cerebral embolism; thrombosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= Sig_peptide
FT   Peptide         19. .28
FT                   /label= Pro_peptide
FT   Protein         29. .273
FT                   /label= Mat_protein
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
```

```
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
DR   N-PSDB; AAV44329.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Disclosure; Page 65; 92pp; English.
XX
CC   This is the deduced amino acid sequence of human mast cell tryptase I
CC   (see also AAV44329). The invention provides: compositions comprising an
CC   isolated tryptase-7 that may include chimeric proteins that contain (a) a
CC   human tryptase for all but the active site region and (b) the substrate-
CC   binding pocket of mouse tryptase-7 or its homologues (see AAW64233-39); a
CC   method for treating a blood clot by administering a nucleic acid molecule
CC   that codes for a tryptase-7, or an expression product, to decrease
CC   fibrinogen activity; a nucleic acid encoding a serine protease (SP); and
CC   a method of producing a mature SP by expressing the inactive zymogen in a
CC   host cell, and cleaving the enterokinase susceptibility domain. The
CC   tryptase-7 polypeptides can be used to treat disorders mediated by
CC   undesirable thrombus clot formation such as myocardial infarction and
CC   reocclusion following angioplasty of blood clots associated with
CC   pulmonary thromboembolism, deep vein thrombosis, cerebral embolism, renal
CC   vein and peripheral arterial thrombosis. They are also useful for all
CC   surgical procedures that require decreased blood clots
XX
SQ   Sequence 273 AA;

  Query Match             92.7%;  Score 1402;  DB 2;  Length 273;
  Best Local Similarity   92.3%;  Pred. No. 1.9e-121;
  Matches  252;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy           3 SLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGS 62
               :||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||||
Db           1 NLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGS 60

Qy          63 LIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADI 122
               |||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||||
Db          61 LIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADI 120

Qy         123 ALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPI 182
               ||||||||||:|| |||| |||||||||||||||||||||||||||| ||||||||||||||
Db         121 ALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPI 180

Qy         183 MENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVV 242
               ||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||||
Db         181 MENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV 240

Qy         243 SWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
               || ||||||||||||||||||||||||||||||
Db         241 SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 273


RESULT 13
AAW63174
ID   AAW63174 standard; protein; 273 AA.
XX
AC   AAW63174;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human mast cell tryptase I polypeptide.
XX
KW   Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
KW   tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
KW   antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
KW   hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
KW   inflammatory skin condition; human; mast cell tryptase I.
XX
OS   Homo sapiens.
```

```
XX
PN   WO9833812-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US001865.
XX
PR   05-FEB-1997;   97US-0037090P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL,  Huang C;
XX
DR   WPI; 1998-437390/37.
DR   N-PSDB; AAV42711.
XX
PT   Tryptase-6 complex inhibitory peptides - used to treat mast cell-mediated
PT   inflammatory disorders e.g. asthma.
XX
PS   Disclosure; Page 46-47; 69pp; English.
XX
CC   This represents the human mast cell tryptase I protein which is a
CC   homologue of the mouse mast cell protease (mMCP-6) zymogen. The invention
CC   provides sequences shown in AAW63160 to AAW63169 that are inhibitors of
CC   mMCP-6. These peptides which are tryptase-6 complex inhibitors, can be
CC   used for treating a mast cell-mediated inflammatory disorder. The
CC   inhibitors can be used to treat inflammatory disorders including asthma,
CC   allergic rhinitis, urticaria and antioedema, eczematous dermatitis
CC   (atopic dermatitis), anaphylaxis, hyperproliferative skin disease, peptic
CC   ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory
CC   skin conditions
XX
SQ   Sequence 273 AA;

  Query Match             92.7%;  Score 1402;  DB 2;  Length 273;
  Best Local Similarity   92.3%;  Pred. No. 1.9e-121;
  Matches  252;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          3 SLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGS 62
              :||||||||||||||||||| ||||: |||||||||||||||||||||   |||||||||
Db          1 NLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGS 60

Qy         63 LIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGADI 122
              |||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||||
Db         61 LIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADI 120

Qy        123 ALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPI 182
              ||||||||||:|| |||| |||||||||||||||||||||||||||| ||||||||||||||
Db        121 ALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPI 180

Qy        183 MENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVV 242
              ||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||||
Db        181 MENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV 240

Qy        243 SWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              || ||||||||||||||||||||||||||||||
Db        241 SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 273


RESULT 14
AAE14348
ID   AAE14348 standard; protein; 691 AA.
XX
AC   AAE14348;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Human protease PRTS-13 protein.
XX
KW   Human; protease; PRTS-13; tranquilliser; gene therapy; vaccine; allergy;
KW   infection; dermatitis; arteriosclerosis; rheumatoid arthritis; hepatitis;
KW   atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; cancer;
KW   gastrointestinal disorder; Cushing's syndrome; seizure; glaucoma; stroke;
```

```
KW   epithelial disorder; urticaria; anorexia; trauma; asthma; eczema; nausea;
KW   hypertension; neurological disorder; Parkinson's disease; drug screening;
KW   cardiant; cell proliferative disorder; multiple sclerosis; osteoporosis;
KW   diabetes mellitus; glomerulonephritis; cardiovascular disorder; anaemia;
KW   autoimmune disorder; inflammatory disorder; myocardial infarction; AIDS;
KW   developmental disorder; reproductive disorder; infertility; diarrhoea;
KW   dementia; acidosis; cataract; gynaecomastia; epilepsy; jaundice.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /label= Signal_peptide
FT   Protein         29. .691
FT                   /note= "Mature human protease PRTS-13 protein"
FT   Active-site     65. .108
FT   Domain          67. .80
FT                   /note= "Type I fibronectin domain"
FT   Active-site     213. .258
FT   Domain          225. .238
FT                   /note= "Type I fibronectin domain"
FT   Domain          371. .389
FT                   /note= "Transmembrane domain"
FT   Active-site     425. .473
FT   Domain          433. .450
FT                   /note= "Kringle domain protein"
FT   Domain          497. .517
FT                   /note= "Transmembrane domain"
FT   Domain          514. .535
FT                   /note= "Kringle domain protein"
FT   Active-site     574. .618
FT   Domain          594. .635
FT                   /note= "Kringle domain protein"
FT   Domain          604. .638
FT                   /note= "Type I fibronectin domain"
XX
PN   WO200183775-A2.
XX
PD   08-NOV-2001.
XX
PF   04-MAY-2001; 2001WO-US014651.
XX
PR   04-MAY-2000; 2000US-0202082P.
PR   11-MAY-2000; 2000US-0203566P.
PR   17-MAY-2000; 2000US-0205803P.
PR   25-MAY-2000; 2000US-0207477P.
PR   01-JUN-2000; 2000US-0209402P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Delegeane AM,  Lal P,  Hafalia A,  Patterson C,  Walia NK,  Kearney L;
PI   Tribouley CM,  Khan FA,  Yao MG,  Baughn MR,  Azimzai Y,  Elliott VS;
PI   Nguyen DB,  Gandhi AR,  Yang J,  Hernandez R,  Policky JL,  Lu DAM;
PI   Reddy R,  Yue H,  Tang YT;
XX
DR   WPI; 2002-034518/04.
DR   N-PSDB; AAD23854.
XX
PT   Novel human proteases and polynucleotides encoding the proteases, useful
PT   for treating, diagnosing or preventing cell proliferative,
PT   cardiovascular, autoimmune/inflammatory, neurological and developmental
PT   disorders.
XX
PS   Claim 1; Page 136-137; 151pp; English.
XX
CC   The invention relates to human proteases (PRTS1-14) and its corresponding
CC   cDNA molecules. Human PRTS and its nucleic acid molecule are useful for
CC   the diagnosis, treatment and prevention of disorders associated with
CC   increased or decreased expression of PRTS. Examples of such disorders
CC   include, cell proliferative disorders (arteriosclerosis, atherosclerosis,
CC   hepatitis, psoriasis and cancers); autoimmune/inflammatory disorders
CC   (AIDS, Addison's disease, allergy, anaemia, asthma, atopic dermatitis,
CC   diabetes mellitus, glomerulonephritis, multiple sclerosis, osteoporosis,
```

```
CC   trauma, Grave's disease, rheumatoid arthritis, ulcerative colitis, and
CC   viral, bacterial, fungal, parasitic, protozoal and helminthic infections)
CC   ; cardiovascular disorders (myocardial infarction, ischaemic heart
CC   disease and hypertension); neurological disorders (epilepsy, Alzheimer's
CC   disease, Pick's disease, Huntington's disease, dementia, Parkinson's
CC   disease, stroke, mental disorders including mood, anxiety and seasonal
CC   affective disorder and prion diseases); gastrointestinal disorders
CC   (Crohn's disease, anorexia, nausea, diarrhoea and jaundice); epithelial
CC   disorders (contact dermatitis, eczema, acne vulgaris, alopecia, scabies,
CC   insect bites and urticaria); reproductive disorder (infertility,
CC   disruption of estrous and menstrual cycle and gynaecomastia); and
CC   developmental disorders (renal tubular acidosis, Cushing's syndrome,
CC   seizure disorders, congenital glaucoma and cataract). PRTS DNA is also in
CC   useful is gene therapy. PRTS and its immunogenic fragments are useful for
CC   screening libraries of compounds in several drug screening assays. The
CC   present sequence is human protease PRTS-13 protein
XX
SQ   Sequence 691 AA;

  Query Match             92.7%;  Score 1401;  DB 5;  Length 691;
  Best Local Similarity   92.0%;  Pred. No. 7.6e-121;
  Matches  253;  Conservative    8;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          1 MLSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 60
              ||:|||||||||||||||||   ||||: |||||||||||||||||||||||||||||||
Db         11 MLNLLLLALPVLASRAYAAPG--QALQRVGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCG 68

Qy         61 GSLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGA 120
              |||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | ||
Db         69 GSLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGA 128

Qy        121 DIALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKV 180
              ||||||||||||:|| |||| |||||||||||||||||||||||||| |||||||| |||
Db        129 DIALLELEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKHVKV 188

Qy        181 PIMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAG 240
              ||||||||||||||||||||||||:|||||||||::||||:|||||||||||||||||||
Db        189 PIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAG 248

Qy        241 VVSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              || | ||||||||||||||||||||||||||||||
Db        249 VVRWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 283


RESULT 15
AAW64240
ID   AAW64240 standard; protein; 274 AA.
XX
AC   AAW64240;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human mast cell tryptase II/beta.
XX
KW   Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;
KW   blood clot; anticoagulant; myocardial infarction; reocclusion;
KW   thromboembolism; cerebral embolism; thrombosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= Sig_peptide
FT   Peptide         19. .28
FT                   /label= Pro_peptide
FT   Protein         29. .274
FT                   /label= Mat_protein
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
DR   N-PSDB; AAV44330.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Disclosure; Page 66-67; 92pp; English.
XX
CC   This is the deduced amino acid sequence of human mast cell tryptase
CC   II/beta (see also AAV44330). The invention provides: compositions
CC   comprising an isolated tryptase-7 that may include chimeric proteins that
CC   contain (a) a human tryptase for all but the active site region and (b)
CC   the substrate-binding pocket of mouse tryptase-7 or its homologues (see
CC   AAW64233-39); a method for treating a blood clot by administering a
CC   nucleic acid molecule that codes for a tryptase-7, or an expression
CC   product, to decrease fibrinogen activity; a nucleic acid encoding a
CC   serine protease (SP); and a method of producing a mature SP by expressing
CC   the inactive zymogen in a host cell, and cleaving the enterokinase
CC   susceptibility domain. The tryptase-7 polypeptides can be used to treat
CC   disorders mediated by undesirable thrombus clot formation such as
CC   myocardial infarction and reocclusion following angioplasty of blood
CC   clots associated with pulmonary thromboembolism, deep vein thrombosis,
CC   cerebral embolism, renal vein and peripheral arterial thrombosis. They
CC   are also useful for all surgical procedures that require decreased blood
CC   clots
XX
SQ   Sequence 274 AA;

  Query Match             92.6%;  Score 1400;  DB 2;  Length 274;
  Best Local Similarity   92.0%;  Pred. No. 3e-121;
  Matches  252;  Conservative    8;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          2 LSLLLLALPVLASRAYAAPAPVQALQQAGIVGGQEAPRSKWPWQVSLRVRDRYWMHFCGG 61
              |:||||||||||||||||||| ||||: |||||||||||||||||||||   ||||||||
Db          1 LNLLLLALPVLASRAYAAPAPGQALQRVGIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGG 60

Qy         62 SLIHPQWVLTAAHCLGPDVKDLATLRVQLREQHLYYQDQLLPVSRIIVHPQFYIIQTGAD 121
              ||||||||||||||:|||||||| |||||||||||||||||||||||||||||  | |||
Db         61 SLIHPQWVLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAD 120

Qy        122 IALLELEEPVNISSRVHTVMLPPASETFPPGMPCWVTGWGDVDNDEPLPPPFPLKQVKVP 181
              |||||||||| :|| |||| |||||||||||||||||||||||||| |||||||||||||
Db        121 IALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVP 180

Qy        182 IMENHICDAKYHLGAYTGDDVRIIRDDMLCAGNSQRDSCKGDSGGPLVCKVNGTWLQAGV 241
              |||||||||||||||||||||||:|||||||||::||||:||||||||||||||||||||
Db        181 IMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGV 240

Qy        242 VSWDEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 275
              ||| ||||||||||||||||||||||||||||||
Db        241 VSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 274

Search completed: August 27, 2005, 23:26:46
Job time : 127 secs
```